Mathematical Modelling of Sound Production in Birds

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biophysik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin

von
Herr Dipl.-Math. Riccardo Zaccarelli

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. h.c. Christoph Markschies

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Dr. Christian Limberg

Gutachter:

1. Prof. Dr. Hanspeter Herzel
2. Prof. Dr. Isao Tokuda
3. Prof. David Berry, PhD

eingereicht am: 1. Oktober 2007
Tag der mündlichen Prüfung: 29. Januar 2008

Abstract

In this thesis, the physics of birds phonation is discussed using a two-mass models approach and the theory of nonlinear dynamics. Two-mass models of the human larynx (rescaled two-mass model and trapezoidal model) have been adapted to the dimension of the avian syrinx to study pressure onset, control of harmonic overtones and "registers" of the sound radiated by the birds vocal organ (syrinx) in the absence of source-tract coupling. Our simulations are a first step towards more realistic modelling of the syrinx. A detailed bifurcation analysis of the trapezoidal model confirms that the geometry and the rest position of the syrinx can influence the harmonic spectra drastically, suggests possible mechanisms involved in the production of rich-harmonic spectra during inspiration and is used to describe quantitatively the contribution of syringeal muscles. The latter is implemented in the model by means of driving time-dependent parameters controlling the labia rest position and frequency modulation. The main focus of the present work is on:

- mathematical modelling of the avian syrinx to study the effects of small dimensions on the pressure onset and the effects of the syrinx geometry and rest position on the harmonic spectra of the sound produced at the vibrating source.

- bifurcation analysis of the trapezoidal model. Results are employed in the implementation of syringeal muscles by means of time-dependent parameters and contribute to explore possible mechanisms involved in the production of more rich harmonic spectra, in particular regarding the inspiratory part of the vocalization of a ring dove (coo).

- model implementation of time-dependent parameters to study quantitatively the contribution of the syringeal muscles on the labia rest areas and control of the fundamental frequency.

In Chapter 3, we analyze two symmetric two-mass models of the avian syrinx. Our first model (rescaled two-mass model) applies to songbirds and is a rescaled version of the well-known human two-mass model. Our second model (trapezoidal model) introduces a smoother geometry and is used to simulate the ring dove (*Streptopelia risoria*) syrinx. Simulations show that both models exhibit self-sustained vibrations and that the intensity of harmonics depends strongly on the configuration of the syrinx. The rescaled two-mass model does not present instabilities. The trapezoidal model, however, displays coexisting limit-cycles that represent vibrations with, and without collisions at the same pressure. Register-like transitions are accompanied by subharmonics and deterministic chaos. In the rescaled two-mass model almost pure tones are found near the onset of vibrations and strong harmonics appear at higher pressures due to collisions. For a small upper mass and a rectangular geometry, collisions leading to strong harmonics can be avoided only near the phonation onset. At higher pressures counteracting forces would be required to diminish collisions. We hypothesize that the avoidance of strong collisions in song birds might be achieved by the medial tympaniform membranes (MTM) that are continuous with the inner vibrating labia. In our model of the ring dove syrinx no collisions occur at default parameters. Consequently, harmonics are fairly weak. The smoother configuration and equal upper and lower masses counteract collisions even at relatively high pressures. This is presumably due to a stronger effect of the subsyringeal pressure acting on both masses. Our simulations reveal that the configuration of the syrinx influences the intensity of overtones. Therefore, the amount of energy in the harmonics could also be controlled by syringeal muscles that directly affect the configuration of the syrinx.

In Chapter 4 we aim to study the contribution of different control parameters in the coo of the ring dove at the syrinx level. The biomechanics of the syrinx, in fact, is very complex and not well understood. The neuromuscular control of vocalisation in birds requires a complicated and precisely coordinated motor control of the vocal organ (i.e. the syrinx), the respiratory system and upper vocal tract. We designed and implemented a quantitative biomechanical syrinx model, i.e., a slightly modified version of the trapezoidal model that is driven by physiological control parameters and includes a muscle model. Our simple nonlinear model reproduces the coo, including the inspiratory note, with remarkable accuracy and suggests once more that harmonic content of song can be controlled by the geometry and rest position of the syrinx. Furthermore, by systematically switching off control parameters, we demonstrate how they affect amplitude and frequency modulation and we generate new experimentally testable hypotheses. Independent control of amplitude and frequency is not possible with the simple syringeal morphology of the ring dove. We speculate that songbirds evolved a syrinx design that uncouples the control of different sound parameters and allows for independent control. This evolutionary key innovation provides an additional explanation for the rapid

diversification and speciation of the songbirds.

Keywords:
syrinx, harmonic overtones, two-mass model, neuromuscular control

Zusammenfassung

In dieser Arbeit wird die Physik der Stimm- und Lautbildung von Vögeln untersucht, wobei zwei verschiedene Zwei-Massen-Modelle des menschlichen Kehlkopfes, sowie die Theorie der Nichtlinearen Dynamik verwendet werden. Die Zwei-Massen-Modelle des menschlichen Kehlkopfes (ein angepasstes Zwei-Massen-Modell und ein Trapez-Modell) wurden dazu an die Größe des Stimmorgans der Vögel (Syrinx) angepasst, um die Druckentstehung, die Steuerung der harmonischen Obertöne und "Register" studieren zu können, die vom Syrinx ohne Kopplung an Quelle und Vokaltrakt erzeugt werden. Unsere Simulationen sind ein erster Schritt in Richtung eines realistischeren Modells der Syrinx. Eine detaillierte Bifurkationsanalyse des Trapezmodells bestätigt, dass die Geometrie und die Ruhelage der Syrinx das harmonische Spektrum drastisch beeinflussen können, und sie gibt Hinweise über mögliche Erzeugungsmechanismen der reichhaltigen harmonischen Spektren während der Einatmung. Des weiteren wird die Bifurkationsanalyse benutzt, um den Beitrag der Muskeln der Syrinx quantitativ zu beschreiben. Dies geschieht in dem Modell mit Hilfe von zeitabhängigen Parametern, welche die Ruhelage der Labia und die Frequenzmodulation steuern. Diese Arbeit hatte folgende Schwerpunkte:

- Mathematische Modellierung der Syrinx von Vögeln, um den Einfluss ihrer Größe auf die Druckentstehung, sowie den Einfluss ihrer Geometrie und Ruhelage auf das harmonische Spektrum der Klangquelle zu untersuchen.

- Bifurkationsanalyse des Trapezmodells. Die Ergebnisse dieser Analyse wurden genutzt, um die Muskeln der Syrinx berücksichtigen zu können. Dies geschieht durch zeitabhängige Parameter variationen. Des weiteren konnte die Bifurkationsanalyse dazu beitragen, Vermutungen über mögliche Erzeugungsmechanismen der reichhaltigen harmonischen Spektren, insbesondere während der Einatmungsphase der Stimmgebung einer Lachtaube (coo), aufzustellen.

- Implementation der zeitabhängigen Parameter in das Modell, um quantitative Aussagen über den Beitrag der Muskeln der Syrinx auf die Ruhepositionen der Labia und die Steuerung der Grundfrequenz zu erhalten.

Im Kapitel 3 haben wir zwei symmetrische Zwei-Massen-Modelle der Syrinx eines Vogels analysiert. Unser erstes Modell kann auf Singvögel angewandt werden, und ist eine an die Größe angepasste Version des bekannten Zwei-Massen-Modells des Menschen. Unser zweites Modell (das Trapezmodell) führt eine geglättete Geometrie ein und wird benutzt, um die Syrinx der Lachtaube (Streptopelia risoria) zu simulieren. Die Simulationen zeigen, dass beide Modelle selbst-erhaltende Vibrationen aufweisen und dass die Intensität der Harmonischen stark von der Konfiguration der Syrinx abhängt. Das erste Modell wies keine Instabilitäten auf. Das Trapezmodell hingegen zeigte gleichzeitig existierende Grenz-zyklen, die Vibrationen mit und ohne Syrinx-schluss darstellen. Register-ähnliche Übergänge werden von Subharmonischen und deterministischem Chaos begleitet. In dem an die Größe angepassten Zwei-Massen-Modell werden nahezu reine Töne in der Nähe des Einsatzpunktes der Vibrationen gefunden. Des weiteren entstehen durch Kollisionen bei höherem Druck starke Harmonische. Verwendet man eine kleine obere Masse und eine rechteckige Geometrie, dann können diese Zusammenstöße, welche starke Harmonische verursachen, nur in der Nähe des Einsatzpunktes der Stimmbildung umgangen werden. Bei höheren Drücken würden entgegenwirkende Kräfte notwendig sein, um diese zu vermindern. Unsere Hypothese ist, dass die starken Zusammenstöße bei Singvögeln durch die “medial tympaniform membranes” (MTM) verhindert werden, welche stetig mit der inneren vibrierenden Labia verbunden sind. In unserem Modell der Syrinx der Lachtaube gab es mit Standardparametern keine Kollisionen. Dadurch sind die Harmonischen sehr schwach ausgeprägt. Die geglättete Konfiguration mit der gleichen oberen und unteren Masse wirkt den Zusammenstößen selbst bei relativ hohen Drücken entgegen. Das ist vermutlich auf einen stärkeren Einfluss des subsyringealen Drucks zurückzuführen, welcher auf beide Massen wirkt. Unsere Simulationen zeigen, dass die Konfiguration der Syrinx die Intensität der Obertöne beeinflusst. Daher kann die Energie in den Harmonischen auch von den syringealen Muskeln gesteuert werden, die direkten Einfluss auf die Konfiguration der Syrinx haben.

In Kapitel 4 ist es unser Ziel, den Einfluss verschiedener Steuerparameter der coo der Lachtaube auf Höhe der Syrinx zu untersuchen. Die Biomechanik der Syrinx ist jedoch sehr komplex und nicht gut verstanden. Die neuromuskuläre Steuerung der Stimmerzeugung in Vögeln setzt eine komplizierte und präzise koordinierte Bewegung des Vokalorgans (Syrinx), des Atmungsorgans und des oberen Vokaltraktes voraus. Wir haben ein quantitatives biomechanisches Syrinx-Modell entwickelt und implementiert. Dieses Modell ist eine leicht modifizierte Version des Trapez-Modells ist, welches durch physiologische Steuerparameter angetrieben wird und zusätzlich ein Modell der Muskeln enthält. Unser einfaches nichtlineares Modell reproduziert die coo, inklusive spezifische Eigenschaften bei der Einatmung, mit erstaunlicher Genauigkeit und legt erneut nahe, dass der harmonische Inhalt eines Gesangs durch die Geometrie und Ruhelage der Syrinx gesteuert wird. Des weiteren zeigen wir, wie die Amplituden- und Frequenzmo-

dulation durch systematisches Ausschalten der Steuerparameter beeinflusst wird, und erzeugen somit neue experimentell überprüfbare Hypothesen. Die unabhängige Steuerung der Amplituden- und Frequenzmodulation ist mit der einfachen syringealen Morphologie der Lachtaube nicht möglich. Wir vermuten, das Singvögel durch die Evolution einen Aufbau der Syrinx entwickelt haben, der verschiedene Steuerparameter der Klangerzeugung entkoppelt und somit deren unabhängige Steuerung ermöglicht. Diese evolutionäre Neuerung bietet eine weitere Erklärung für die schnelle Diversifikation und Spezifikation von Singvögeln.

Schlagwörter:
Syrinx, Oberwelle, Zwei-Masse Modell, neuromuskuläre Kontrolle

Contents

List of Figures

List of Tables

1

Introduction

The evocative power of bird songs represents an extremely interesting problem for the scientific community. Among birds, songbirds (Oscines), parrots and hummingbirds share with humans the remarkable feature of learning their vocalisations after a certain degree of exposure to a tutor. For these reasons, the possible existence of common principles of learning and memory underlying human speech and birdsong has made Oscines an excellent and widely used model system for sensorimotor learning and human speech acquisition (Doupe and Kuhl, 1999). The control of the fundamental frequency of the sound and the gating of sound elements happens at the level of the syrinx (Greenewalt, 1968; Goller and Suthers, 1996a,b; Elemans et al., 2004). By illuminating the mechanisms by which the activities of different syringeal muscles control the avian vocal organ for song production, several models have been developed to build a bridge between neural activity and the physical manifestation of the motor control, i.e., the song (Laje et al., 2002; Laje and Mindlin, 2002; Mindlin et al., 2003; Laje and Mindlin, 2005; Mindlin and Laje, 2005).

However, the biomechanics and neuromuscular control of the syrinx, which in birds is located at the bifurcation of the trachea into the bronchi, is very complex and not well understood. An emergent model organism to study the biomechanics of phonation and its control is the ring dove (*Streptopelia risoria*). The syrinx morphology of this species is relatively simple with only two paired muscles controlling its geometry, compared to 6-8 pairs in most songbirds (King, 1989). Because this non-songbird does not learn the syntax of its species-specific song as songbirds do (Nottebohm, 1972), it is of less interest to study learning. However, juveniles of a closely related species still require motor practice to utter the final characteristic coo (Ballintijn and ten Cate, 1997), which indicates that the individuals must go through numerous iterations to match their vocal output with some sort of neural template.

Let alone vocal learning, the mechanisms underlying the great variety of birds vocal utterances, the extremely complex structure of some songs and the quality of some sounds represent a further interesting scientific problem. Only in last years the understandings of mechanisms underlying sound production in birds advanced very rapidly (Elemans et al., 2003). Despite the existence of several qualitative models to explain sound generation in birds (Greenewalt, 1968; Gaunt and Gaunt, 1985; Goller and Larsen, 2002), until recent years there were very few attempts, compared to the situation for the human larynx and vocal tract, to analyze the avian vocal organ, the syrinx, in quantitative terms (Brackenbury, 1979; Fletcher, 1988). For decades, the most widely accepted mechanism of sound production in birds (*classical model*) identified an edge-clamped membrane present in the syrinx (the MTM) as the vibrating structure, while a soft tissue mass, the lateral labia (LL) were thought to provide constriction of the syringeal lumen (Greenewalt, 1968; Casey and Gaunt, 1985; Gaunt and Gaunt, 1985, and references therein). More recently, Goller and Larsen (1997a,b) demonstrated that labia (medial labia, ML, in songbirds, lateral tympaniform membranes, LTM, in pigeons), rather than thin membranes, are the source of sounds in birds. Mathematical modelling of the avian syrinx has received a new impulse due to these findings, suggesting that sound generation mechanism in birds could be similar to that operating in human phonation (labial model). In this new perspective, the syrinx system is amenable to nonlinear dynamics analysis, which can provide an entirely different perspective on the underlying basis of phonatory function (Titze et al., 1993). Human vocal production demonstrates that certain irregular phenomena seen in human pathological voices and baby crying result from nonlinearities in the vocal production system (Mende et al., 1990; Herzel and Wendler, 1991; Titze et al., 1993; Herzel, 1993; Herzel et al., 1994), and equivalent phenomena are quite common in nonhuman primate vocal repertoires (Wilden et al., 1998; Fitch et al., 2002; Tokuda et al., 2002). In birds, it has been shown that, rather than direct muscular manipulation, the origin song features (frequency-mode locking and register transitions) in the zebra finch (*Taeniopygia guttata*) are due to intrinsic (nonlinear) dynamics in the syrinx. As a check consistency, a model based on two-mass models of human phonation (Ishizaka and Flanagan, 1972) has been developed (Fee et al., 1998; Fee, 2002). Songbirds syrinx is a complex bipartite structure whose two sides are potentially capable of acting either independently or together to produce sound (Suthers, 1990; Suthers et al., 1999). Laje and Mindlin (2005) proposed a model to study the acoustic interactions between their two sound sources, as well as the acoustic coupling between sources and vocal tract, suggesting that there could be mechanisms, other than intrinsic nonlinearity of the labia, underlying complex features in the radiated sound.

Moreover, the ability of some bird species to produce pure tonal sounds has always turned into an interesting scientific problem which is still widely debated (Klatt and Stefanski, 1974; Nowicki, 1987; Williams et al., 1989; Nowicki et al., 1992; Beckers et al.,

2003a, 2004; Riede et al., 2004; Fletcher et al., 2004): some species have whistle-like songs and not much energy is found in the harmonics (Nowicki et al., 1992), whereas other species, such as the zebra finch (*Taeniopygia guttata*), display strong harmonics during song or calls. In this framework, theoretical difficulties of the classical model in explaining how a vibration-based mechanism can generate pure tonal sounds led to alternative hypothesis, e.g., vortex shedding through a fixed constriction of the syrinx (whistle model) with no vibrations involved, in a similar manner to that operating in a hole-tone whistle (Gaunt et al., 1982; Casey and Gaunt, 1985). However, experimental evidence by means of songs recorded in heliox atmosphere (which speeds up sound velocity, and consequently, in case of a hole-tone whistle mechanisms, should also shift upwards the fundamental frequency) did not induce substantial shifts of fundamental frequency in any of the investigated species (songbirds: Nowicki, 1987; pigeons: Ballintijn and ten Cate, 1998; parrots: Brittan-Powell et al., 1997b). These findings provide evidence for a vibration-based mechanism, but do not make it possible to discern between the classical model (membrane vibration hypothesis) and alternative models such as the labial model. In this case, the most compelling evidence which undermines the assumption of vibrating membranes came from Goller and Larsen (1997a,b).

Therefore, there are currently three hypothesis to explain possible mechanisms underlying the production of pure tonal sounds (Beckers et al., 2003a): 1) a vibrating valve produces a multifrequency harmonic source sound which is linearly filtered by trachea resonances (source-filter model) widely used to model the human vocal tract (Fant, 1960), 2) vocal tract resonances are coupled to a vibrating valve and suppressing the normal production of harmonics at the source (analogous to human soprano singing), 3) pure-tone sound is produced by a yet-unknown source mechanism. In ring doves (*Streptopelia risoria*) it has been shown that there is direct evidence to support the source-filter model for the suppression of higher harmonics (Beckers et al., 2003a), and that the combined influence of the trachea, glottis and inflated upper esophagus acts as an effective band-pass filter to eliminate higher harmonics generated by the syrinx (Riede et al., 2004; Fletcher et al., 2004). In songbirds, the inaccessibility of the syrinx and the relatively small size of songbirds make experimental tests more difficult. Moreover, because of the complex morphology of the oscine syrinx, several hypothesis still hold, and several mechanisms can not be ruled out entirely, even aerodynamic processes such as the whistle hypothesis for some bird species (Ballintijn and ten Cate, 1998; Goller and Larsen, 2002; Beckers et al., 2003a). Recently, Riede et al. (2006) have shown how songbirds seem to tune their vocal tract to match the fundamental frequency of the song. However, it is not necessary to invoke source mechanisms radically different from valves vibrating largely independently from resonant cavities such a trachea and beak, to explain pure-tonality (Beckers et al., 2003a).

In this framework, this text deals with modelling of the avian vocal organ, the syrinx,

with particular attention to the syrinx of the ring dove (*Streptopelia risoria*) and the associated superfast syringeal muscles affecting syrinx variations in geometry and tension (Elemans et al., 2004, 2006). We will first adapt two modified versions of the well-known two-mass model (widely used to study vocal fold vibrations) to the dimensions of the bird syrinx. These core models allow comprehensive bifurcation analysis and we can address basic questions: Do we find reasonable vibrations of the masses even for much smaller geometries? Can a modified geometry (adapted to the syrinx anatomy) influence voice onset and harmonic spectra? Furthermore, we will incorporate the dynamic control of associated superfast syringeal muscles in the trapezoidal model. This will allow a quantitative comparison of observed bird songs and simulations.

In Chapter 2, the mechanism of sound production in humans based on the myo-elastic aerodynamic theory of phonation (van den Berg et al., 1957) will be reviewed. The myo-elastic aerodynamic theory implies that phonation can be regarded as a biomechanical aerodynamical system. Biomechanics and aerodynamics employed in vocal folds modelling will be reviewed, and we will describe the two-mass models of human phonation which will be recalled in the models of the avian syrinx. As a bridge between humans and birds, an introduction to the theory of nonlinear dynamics will be given and will be related to phonation and animal vocalization. The final section will describe the anatomy of birds syrinx, the mechanisms of sound production in birds, and the avian models developed in the last years, divided into classical and modern, relative to the findings of Goller and Larsen (1997a,b).

In Chapter 3, we will analyze two symmetric two-mass models of the avian syrinx. Our first model (rescaled two-mass model) applies to songbirds and is a rescaled version of the well-known human two-mass model (Ishizaka and Flanagan, 1972; Steinecke and Herzel, 1995). Our second model (trapezoidal model) introduces a smoother geometry as in Pelorson et al. (1994) and Lous et al. (1998), and is used to simulate the ring dove syrinx. A detailed description of the geometry and the implementation of the forces is given in Appendix A. Our simulations will show that rescaled biomechanical models originally developed to describe mammalian vocal fold vibrations can be adapted to model the bird syrinx.

In the songbirds model (rescaled two-mass model), collisions leading to strong harmonics can be avoided only near the phonation onset. At higher pressures, counteracting forces would be required to diminish collisions. In songbirds, the role of the MTM has been widely speculated: it may play a distinct role especially for the generation of high-frequency sounds (Goller and Larsen, 2002), or contribute to the wave-like movements of the MVM (MTM +ML) resulting in a more efficient sound generation than possible with a one-mass oscillator (Titze, 1988, 1994; Fee, 2002). In this framework, we will

hypothesize that the avoidance of strong collisions in song birds might be achieved by the MTM that are continuous with the inner vibrating labia (Fee, 2002). Contrarily, in our model of the ring dove syrinx (the trapezoidal model) no collisions occur at default parameters. Consequently, harmonics are fairly weak. Our simulations reveal that the configuration of the syrinx influences the intensity of overtones. Therefore, the amount of energy in the harmonics could also be controlled by syringeal muscles that directly affect the configuration of the syrinx. Furthermore, the syrinx is in a persistent convergent shape. As such, the suprasyringeal pressure has little effect on the masses, which might lead to reduced source-tract interaction. This again suggests as in Beckers et al. (2003a), that filter mechanisms in birds are closer to those in human phonation (e.g., Fant, 1960) in contrast to other possible explanations, such as the soprano-model.

In the rescaled two-mass model we found no voice instabilities. On the other hand, we will show that our trapezoidal model exhibits coexistence of attractors and jumps from harmonic rich spectra to more sinusoidal oscillations of the radiated sound pressure. Ring doves already exhibit stronger harmonics during inspiratory phonation compared to expiratory phonation even at low intensities (Gaunt et al., 1982; Beckers et al., 2003a; Elemans, 2004; Riede et al., 2004). This implies that asymmetries between outflow and inflow of the air have to be taken into account. We will show in Appendix B and in Chapter 4 that asymmetries in the prephonatory syringeal shape and different prephonatory syringeal widths could be involved in the generation of more rich dynamics. Instabilities such as frequency jumps are commonly observed in the ring dove coo (Beckers et al., 2003b; Elemans, 2004). However, although intrinsic nonlinear properties of the syrinx add complexity to the level of motor control (Fee et al., 1998; Fee, 2002; Zaccarelli et al., 2006), only muscle control can explain the fast but gradual frequency modulations of ring dove coos (Elemans et al., 2004).

In Chapter 4, we will incorporate the dynamic control of associated superfast syringeal muscles in the trapezoidal model (for the derivation of the time dependent parameters describing the syringeal muscles activity, see Appendix C). The quantitative approach will confirm the validity of two hypothesis: first, in ring doves the tracheolateralis muscle (TL) modulates the position of the Lateral Vibratory Masses (LVM) and as such the syringeal aperture. Second, both TL stress and the transmural pressure (pressure difference between the air sac surrounding the syrinx - the interclavicular air sac - and syringeal lumen) affect tension in the LVM. We will present the first model of bird song that implements a muscle model to calculate force produced by the muscle. In a typical ring dove coo, register transitions to aphonia, period doublings or chaos are rare during expiration or inspiration. The simulations with physiological input parameters will also not exhibit transitions to different dynamical regimes. Moreover, the simulated oscillation maximal amplitude will be consistent with the oscillation amplitude

observed in pigeons (Larsen and Goller, 1999). Therefore, our simple nonlinear model reproduces the coo with great accuracy and as such will provide strong evidence for our biomechanical framework. This is not a trivial result, because in nonlinear dynamical systems different excitations may result in different time-courses, described by nonlinear dynamics theory (Guckenheimer and Holmes, 1983; Berge et al., 1986; Jackson, 1989; Arrowsmith and Place, 1990; Strogatz, 1994). Moreover, we will discuss testable predictions of in vivo experiments by switching on and off the TL and transmural pressure activity.

Concluding, our simulations will suggest that the control of frequency modulation during sound production is relatively independent of a pressure difference between air sacs compared to direct muscular control in ring doves. Furthermore, the complex interplay of bronchial pressure, syringeal muscle force generation and transmural pressure all affect precise amplitude and frequency modulation. In the case of ring doves, the syringeal muscles have the most profound effect on both amplitude modulation (AM) and frequency modulation (FM). Nevertheless, a separated physiological correlate with AM and FM (Gaunt et al., 1982; Beckers et al., 2003b) seems not to be the case in ring doves. Therefore, by showing that independent control of song characteristics is not possible with the simple syringeal morphology of the ring dove, we speculate that songbirds evolved a syrinx design that allowed for the independent control of sound parameters. Uncoupling the control of different sound parameters as an evolutionary key innovation would lead to a tremendously increase in the possible variation of song, which provides an additional explanation for the rapid diversification and speciation of songbirds.

2

Physics of Sound Production

2.1 Voice production in humans

Physically, sound is generated whenever there is a disturbance of the equilibrium of density (or pressure) of a gas, liquid or solid (Titze, 1994). In humans the primary source of sound for *speech* production is the time-varying glottal airflow, produced by *vocal fold* vibration. The physical and physiological processes of vocal fold vibration is termed *phonation* (Titze, 1994).

The basic anatomy of the human airway system involved in sound production is presented in Fig. 2.1: vocal folds are housed by the *larynx*, an organ placed in the neck of mammals, and they are composed of layered muscular and nonmuscular tissues (Titze, 1994, and references therein). The air space between the vocal folds is called *glottis*. The glottis is an important landmark in speech anatomy: widely used terms are, e.g., *supraglottal* ("above the glottis") and *subglottal* ("below the glottis"). The vocal folds can be regarded as a bilateral, three-dimensional and multilayered anisotropic (not invariant with respect to direction) glottal valve structure.

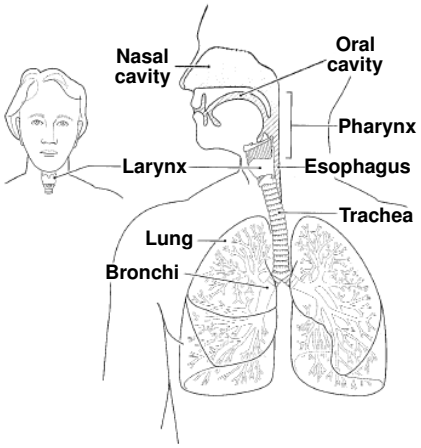


Figure 2.1: Basic components of the airway system in humans (modified after Titze, 1994).

The tract between the glottis and the mouth is called *vocal tract*, and it is highly in-

fluential in determining the vocal sounds produced (see Section 2.1.4). The tract below the larynx and the lungs is composed of the *trachea*, a tube which bifurcates into two *bronchi*, which are eventually connected to the lungs.

Before self-sustained oscillations of the vocal folds set in, activity of adduction muscles (intrinsic muscles) close the glottis. The prephonatory narrowing or closing (*adduction*) of the glottis corresponds to the prephonatory rest position. The lungs contract, pushing air from the lungs toward the mouth. Physically, phonation results from the interplay of a driving fluid (the air) with a complex visco-elastic medium (the vocal folds). Vocal fold oscillations occur due to a flow-induced instability of the visco-elastic vocal fold tissue. The mechanism for self-sustained vocal fold oscillations will be now described in some more details. For a comprehensive review, see Titze (1994, 2002).

2.1.1 Biomechanical oscillators

Biomechanics is the study of motion of living material under known or assumed forces, following the principles of ordinary mechanics: Newton laws of motion, conservation of mass, momentum and energy (Titze, 1994).

One approach widely used for the characterization of vocal fold properties is the reduction of the anisotropic structure of the fold tissue into a small number of condensed elements such as masses and springs with given stiffness (Kob, 2002). This idea is a characteristic aspect of common two-mass models, whereby (self-sustained) mass-spring oscillators can be used to approximate the vocal folds. Mass-spring oscillators are systems made by one or more masses m connected to a rigid foundation by means of a spring. The forces present in the system are inertia, friction of the moving mass, elastic force due to the spring and additionally, in self-sustained mass-spring oscillators, driving forces. In vocal fold modelling the driving forces are typically represented by the aerodynamic forces generated by the lung pressure. The fundamental equation of a self-sustained mass-spring oscillator results:

$$\begin{array}{ccccccc} m\ddot{x} & + & r\dot{x} & + & kx & = & f(x,\dot{x},t) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \text{inertia} & & \text{friction} & & \text{elasticity} & & \text{driving force} \end{array} \tag{2.1}$$

where m is the mass (the inertial property of the mechanical system), k is the *spring constant* which represents the effective stiffness (elasticity) of various tissue layers of the vocal cords and r is the *damping*. As applied to the vocal folds, this element represents the viscosity of the tissue, i.e., the energy absorber in the tissue (Titze, 1994).

The frequency of oscillation of the system is defined as the number of back-and-forth movements that are made per second. For unforced oscillation (i.e., f =0 in eq.

2.1) of a mass-spring system, this frequency is

$$\mathscr{F}_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{2.2}$$

(or alternatively $\omega = \sqrt{\frac{k}{m}}$, where ω is the angular frequency defined as $\omega = 2\pi\mathscr{F}$). Eq. 2.2 suggests that the stiffness-to-mass ratio is a key quantity for the determination of $\mathscr{F}_0$ (Titze, 1994).

Vibration of the vocal folds is said to be self-sustained. Self-sustained oscillations are those in which the energy loss per cycle is zero:

$$\int_{t_0}^{t_0+T} \frac{\mathrm{d}E}{\mathrm{d}t}\,\mathrm{d}t = 0 \tag{2.3}$$

where T is the period of oscillation of the mass and

$$\frac{\mathrm{d}E}{\mathrm{d}t} = (f - r\dot{x})\dot{x} \tag{2.4}$$

is the rate of change of energy with respect to time t[1]. If $\frac{\mathrm{d}E}{\mathrm{d}t} < 0$, the oscillation is said to be *damped* (diminish the back and forth movement until the vocal folds come to rest). As observed by Titze (1988), the dependence of f on $\dot{x}$ is crucial for oscillation: whenever f is in the direction of the velocity, energy is imparted to the oscillating mass. Conversely, energy is taken out of the mass-spring system. In vocal folds, under specific flow conditions, an energy transfer from the glottal airstream to the tissue overcomes frictional energy losses and oscillation is set on. We analyze this mechanism in the next section.

2.1.2 Aerodynamics of phonation

Air flow through the glottis involves the study of a fluid confined in a region of space such as tubes or ducts. The term *flow* indicates the volume of fluid passing a given cross section of the tube or duct per second (Titze, 1994), and is measured in $\frac{\mathrm{m}^3}{\mathrm{sec}}$. For an ideal fluid flowing at a steady rate the flow is said to be *steady* or *stationary*. If the fluid density does not change when forced through a constriction of the duct, the flow is said to be *incompressible*.

To understand how aerodynamic energy is converted to acoustic energy, it is necessary to introduce some of the basic laws that govern fluid flow in ducts. Among them are the energy law (Bernoulli law) and the continuity law.

[1] By definition, in a self-sustained mass-spring oscillator system (Berge et al., 1986), $E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2$. Consequently:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = m\ddot{x}\dot{x} + k\dot{x}x = (m\ddot{x} + kx)\dot{x} \stackrel{*}{=} (f - r\dot{x})\dot{x}$$

where in * we used the relation $m\ddot{x} + kx = f - r\dot{x}$ (eq. 2.1).

For an ideal fluid flowing at a steady rate, if no energy losses are encountered during the path through the pipe, Bernoulli energy law reads:

$$P + \frac{\rho v^2}{2} = \text{constant} \tag{2.5}$$

where ρ is the fluid density and v is the particle velocity.

Continuity law (also known as continuity equation or conservation of mass of Navier-Stokes equations) of incompressible fluid confined in a tube or duct states that the air flow at every point along the duct is constant, i.e.

$$vA = U = \text{constant} \tag{2.6}$$

Most vocal fold flow models are based on the stationary Bernoulli equation. Results of investigations on steady pressure-flow profiles (e.g., van den Berg et al., 1957; Ishizaka and Matsudaira, 1972; Titze, 1988, and references therein) and simulations of the Navier-Stokes equation (Liljenkrants, 1991; Iijima et al., 1992) indicate that the flow at glottal entry and throughout the glottis is essentially nonturbulent and describable by a modified Bernoulli equation (several further studies have been performed to gain insight into airflow behavior and intraglottal pressure distribution. See, e.g., Pelorson et al., 1994, and references therein).

Classical descriptions of vocal fold vibration (van den Berg et al., 1957; van den Berg, 1958) usually begin with the assertion that the vocal folds are sucked together by a negative Bernoulli pressure in the glottis. This theory was called *myoelastic-aerodynamic theory of vocal fold vibration* (van den Berg, 1958) and can be summarized in three points which act in a loop:

① Bernoulli pressure sucks vocal folds together.

② collapse of vocal folds is then followed by a buildup of the subglottal pressure which moves them laterally (outward)

③ Elastic forces of the tissue retard the lateral motion and eventually reverse it.

However, as discussed in Titze (1988, 1994), more features are involved in vocal fold vibration than Bernoulli pressure alone. What permits the driving force (determined by the Bernoulli pressure) to lower (or being less positive) during closing than opening? The buildup of subglottal pressure after vocal folds collapse is not sufficient alone as a driving force, because oscillations can occur also without collision (Ishizaka and Matsudaira, 1972; Titze and Scherer, 1983; Titze, 1988, 1994, and references therein). Therefore, the system needs to change the effective driving force alternatively (in particular, at alternate quarter-cycles of oscillations). This driving-force asymmetry can be achieved in at least two different ways (Titze, 1988):

① By making use of oppositely phased subglottal and supraglottal acoustic pressures

② by varying the glottal geometry to create different intraglottal pressure distributions

These two cases are not mutually exclusive and can both occur in normal phonation.

It has been shown (Titze and Scherer, 1983; Titze, 1988, 1994) that the *mean intraglottal pressure* $\overline{P}$ derived from Bernoulli equation can be approximated (for no vocal fold collision and idealized flow in the glottis) by:

$$\overline{P} \simeq P_{\text{sup}} + (P_s - P_{\text{sup}}) \left(1 - \frac{a_2}{a_1}\right) \tag{2.7}$$

where the following quantities (which will be extensively recalled throughout the dissertation) are defined:

- P_s is the *sub*-glottal pressure
- P_{sup} is the *supra*-glottal pressure (input pressure to the vocal tract)
- a_1 is the cross-sectional area of the glottis at glottal entry (inferior lip)
- a_2 is the cross-sectional area of the glottis at glottal exit (superior lip)

Eq. 2.7 predicts that the vocal folds are partially driven by the transglottal pressure $P_t = P_s - P_{\text{sup}}$ and partially by the vocal tract input pressure P_{sup}.

Flanagan and Landgraf (1968) proposed a one-mass model of the vocal folds. This model predicts an acoustic coupling of the glottis to the vocal tract. This condition must be met because a one-mass model cannot take into account glottal geometry, and therefore from eq. 2.7, $\overline{P} = P_{\text{sup}}$ because $a_1 = a_2$. This almost trivial relation suggests that something must happen above the glottis to change this pressure during the glottal cycle. The key element is the inertia of the air in the vocal tract. As a result of air inertia (see Titze, 1988, 1994 for a detailed explanation), the shape of the airflow during open glottis is characterized by a slow rise and an abrupt fall with respect to the shape of the tissue displacement (see Fig. 2.2). Consequently, the shape of the air particle velocity $v = \frac{U}{A}$ displays a gradual increase over the open portion of the cycle, resulting in a mean intraglottal pressure $\overline{P} = P_s - \frac{1}{2}\rho v^2$ (and therefore,

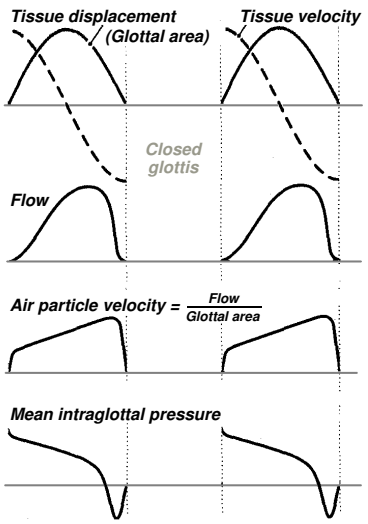


Figure 2.2: Waveforms during sustained oscillation in the open glottis (modified after Titze, 1994).

a net driving force) less positive during closing than opening (see Fig. 2.2). Rothenberg (1981) has showed that at frequencies below the first formant frequency (the usual case for speech) the air column in the vocal tract is primarily inertive, i.e. it acts like a mass of air that is accelerated and decelerated as a unit (Titze, 1988). Inertive vocal tract can be expressed by the relation:

$$P_{\text{sup}} = L\frac{\mathrm{d}U}{\mathrm{d}t} \tag{2.8}$$

which derives from the second law of motion (compare momentum $F = m\frac{\mathrm{d}v}{\mathrm{d}t}$) with the mass "replaced" by the quantity L (*inertance*). Concluding, the Bernoulli pressure alone is not sensitive to the direction (inward versus outward) of vocal fold movement; hence, it cannot by itself supply the necessary velocity-dependent force to transfer energy from the fluid to the tissue. Only in conjunction with vocal tract inertance can the proper asymmetries between driving force and tissue velocity be established.

One-mass models as Flanagan and Landgraf (1968) predicted the above mentioned acoustic coupling of the glottis to the vocal tract. Two-mass models of vocal fold vibration, however, can take into account the tissue displacement. Excised larynx experiments show that even without a vocal tract similar vibratory pattern of the vocal folds can be observed (Titze, 1994; Herzel and Knudsen, 1995, and references therein). Thus, it is justified to neglect the interaction of the glottal flow with sub- and supraglottal resonances as a first approximation. It has been shown (Scherer and Titze, 1983) that $\overline{P}$ is larger for a convergent shape than for a divergent shape: if we assume for simplicity that the vocal tract is disconnected, i.e. $P_{\text{sup}} = 0$, we see from eq. 2.7 that

- convergent glottis $(a_1 > a_2) \Leftrightarrow \overline{P} > 0$
- divergent glottis $(a_2 > a_1) \Leftrightarrow \overline{P} < 0$

In particular it has been shown (Ishizaka and Matsudaira, 1972) that Bernoulli equation is less applicable to the divergent glottis (Titze, 1994): the air flow often detaches from the surface of the vocal folds and forms a jet which keeps the pressure in the glottis close to zero (*jet separation assumption*). However, the important aspect is that the pressure is less positive during closure than during opening: energy is transferred from the airstream to the tissue because the net driving force over the cycle is synchronized with tissue movement (see Fig. 2.3). Nonuniform tissue movement is best described in terms of normal modes (vibratory patterns) of vibration. The number of normal modes equals the number of degrees of freedom of the system. Titze (1976) has calculated the normal modes in the model of Ishizaka and Flanagan (1972) obtaining the normal modes frequencies $(\omega_1, \omega_2) = (120, 201)$ Hz. In the lower mode the masses are moving in phase and have almost the same amplitude, whereas in the higher mode the masses move 180° out of phase and the smaller mass has much larger amplitude. Whether one,

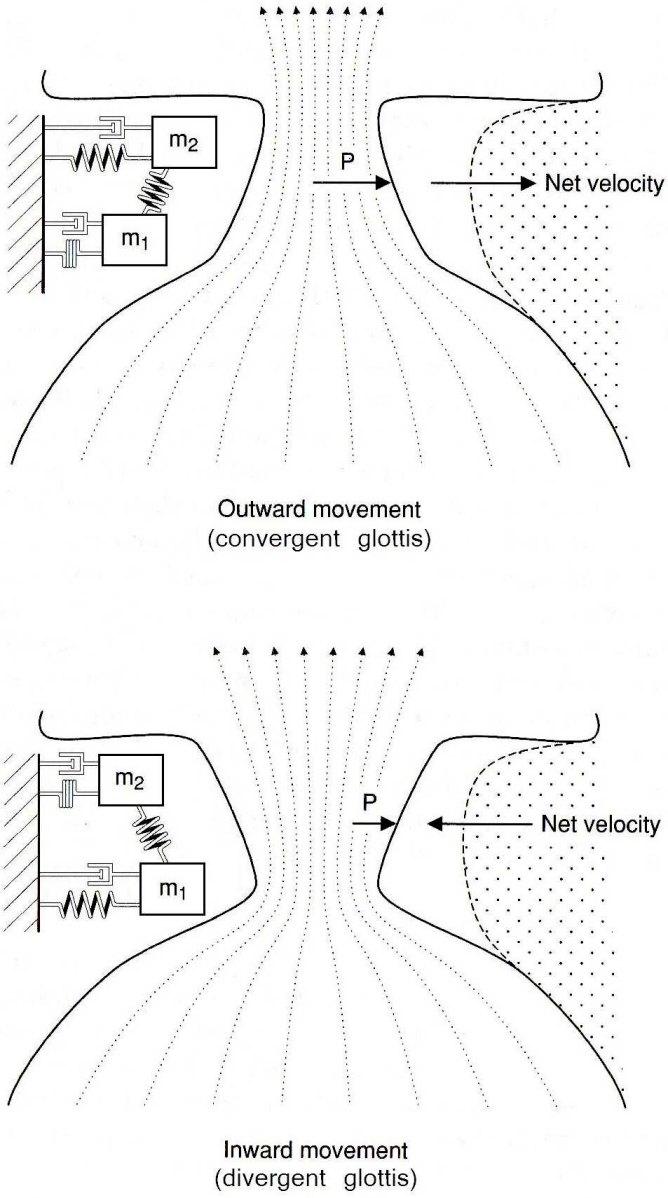


Figure 2.3: Sketch of the classical two-mass model (modified after Titze, 1994). Outward movement with convergent glottis (upper panel) and (lower panel) inward movement with a divergent glottis. Note the comparative lengths of the arrows labeled P to indicate that $\overline{P}$ is larger for a convergent shape than for a divergent shape (Titze, 1994).

the other, ore both modes are present depends on the type of excitation.

The one-mass model by Flanagan and Landgraf (1968) reproduced some glottal features and was used successfully for the synthesis of voice and voiceless sound (Ishizaka and Flanagan, 1972). However, repeated observation of the vocal folds in slow motion has demonstrated that the vocal folds do not move like a solid bar mass, i.e., uniform tissue displacement *seldom, if ever*, occurs (Titze, 1994). Theoretical work has shown that

a two-mass approximation can account for most of the relevant glottal detail, including phase differences of upper and lower edges (Ishizaka and Flanagan, 1972).

2.1.3 Two-mass models of vocal folds vibration

Based on the principles introduced in sections 2.1.1 and 2.1.2, two-mass models of human vocal fold vibration have been used successfully to describe the normal voice (Ishizaka and Flanagan, 1972; Pelorson et al., 1994), vocal fold paralysis (Isshiki et al., 1978; Smith et al., 1992; Steinecke and Herzel, 1995; Mergell et al., 2000), phonation onset (Mergell et al., 1998), voice instabilities at high pressures (Jiang et al., 2001) and source-tract coupling (Mergell and Herzel, 1997; Hatzikirou et al., 2006). We review in this section few significant works which will be often recalled throughout the text.

Ishizaka and Flanagan (1972)

The "mile stone" of two-mass models is certainly the two mass model of Ishizaka and Flanagan (1972). The model represents the vocal folds as two masses m_1, m_2 coupled together. In the frontal plane[2], x_{0i} is the distance of mass m_i from the glottis midline *at rest*, and consequently $a_{0i} = 2\ell x_{0i}$ represent the rest areas (ℓ usually denotes the length of the glottis on the *transverse* plane). The variation of the glottal areas reads $a_i = a_i(x_i) = 2\ell x_i + a_{0i}$, where x_i are given by the equations of motion

$$m_1\ddot{x}_1 + r_1\dot{x}_1 + s_1(x_1) + k_c(x_1 - x_2) = f_1 \tag{2.9}$$

$$m_2\ddot{x}_2 + r_2\dot{x}_2 + s_2(x_2) + k_c(x_2 - x_1) = f_2 \tag{2.10}$$

and $s_i(x_i) = k_i x_i(1 + \eta_{k_i} x_i^2)$ is a (modified) elastic force which describes the nonlinearity of the spring (as measured in fresh excised human vocal cords) by means of the coefficient η_{k_i} (Ishizaka and Flanagan, 1972).

Collision forces Because a contact force at collision will cause some deformation in the flesh of the vocal folds, this deformation will produce a restoring force which is modelled as an additional elastic force. Therefore, during collision of mass m_i, $s_i = k_i x_i(1 + \eta_{k_i} x_i^2) + C_i(x_i)$, where

$$C_i(x_i) = h_i(x_i + x_{0i})\left(1 + \eta_{h_i}(x_i + x_{0i})^2\right) \tag{2.11}$$

[2]In the literature, the following cross sections are usually defined with respect to the human body:

plane	x axis direction	y axis direction
Frontal	right-left	bottom-top
Transverse	right-left	back-front
Sagittal	back-front	bottom-top

is the force required to produce the deformation to mass m_i, $h_i = 3k_i$ is a linear stiffness and consequently (compare η_{k_i}), η_{h_i} is a coefficient which represents the nonlinearity of the contacting vocal folds. Mathematically, C_i is a nonlinear function depending on how much the mass m_i crosses the midline of the glottis toward the opposite mass.

Pressure forces The forces F_i are obtained by means of the glottal pressure over the labia surface, $F_i = P_i \ell d_i$, where P_i is the pressure acting on the i$^{\text{th}}$ mass. $(P_1, P_2) = (P_s, 0)$ if the masses m_1 collide ($a_1 \leq 0$), $(P_1, P_2) = (P_s, P_s)$ if *only* $a_2 \leq 0$. Otherwise, in case of an *open* glottis (i.e., $a_1, a_2 > 0$) the glottal flow is assumed to be quasi-steady and the pressure distribution along the glottal flow is obtained by means of Bernoulli equation. In account of that, P_1, P_2 read:

$$P_1 = P_s - P^{\text{vena}} - \frac{1}{2}\left(P_1^{\text{visc}} + P_1^{\text{inert}}\right) \tag{2.12}$$

$$P_2 = P_1 - P^{\text{junc}} - \frac{1}{2}\left(P_1^{\text{visc}} + P_1^{\text{inert}}\right) - \frac{1}{2}\left(P_2^{\text{visc}} + P_2^{\text{inert}}\right) \tag{2.13}$$

As we see, several corrections are added to the Bernoulli equation. In particular:

- P^{vena} (*vena contracta* effect) is a modified Bernoulli pressure due to the abrupt contraction in cross-sectional area at the inlet of the glottis which produces a vena contracta surrounded by stagnant air. The vena contracta makes the area a_1 appear smaller than it actually is and the pressure drop greater than that dictated (Bernoulli law) by an ideal area change.

- P_i^{visc} is a pressure drop governed by viscous losses. Throughout the constrictions formed by the lower and upper cord edges m_i, the pressure falls linearly according to a resistance to the volume flow proportional to $\frac{d_i}{a_i^3}$.

- P_i^{inert} is an additional term induced by the inertance of air, $L_i = \frac{\rho d_i}{a_i}$

- P^{junc} is a pressure drop due to the abrupt change of the glottal cross-sectional area in the junction between masses 1 and 2.

Finally, at the outlet of the glottis the pressure recovers toward the atmospheric pressure. The pressure recovery, based on quasi-steady momentum conservation equation, does not affect the motion of the masses but is significant to the acoustic interaction between the vocal tract and the cord source (Flanagan and Landgraf, 1968). As the supraglottal duct area is much larger than the minimum cross-sectional area within the glottis, the pressure recovery is very small (Pelorson et al., 1994; Story and Titze, 1995).

Pelorson et al. (1994)

In this work a theoretical model is presented with an improved aerodynamic description of glottal flow which takes into account the formation of a free jet downstream of a moving separation point in the closing phase of the glottal cycle (Sciamarella and D'Alessandro, 2004). The theoretical model (based on the boundary-layer approximation), as well as the assumptions made, are validated using steady- and unsteady-flow measurements combined with flow visualization (Pelorson et al., 1994). In order to evaluate the effective impact of the revised theory, an application to a novel two-mass model is presented. The model considers a two-point parametric curve describing the shape of the vocal folds. In other words, the masses are not anymore rectangular, but cylinders connected by tangent plates: in fact, it would not make much sense to introduce a quasi-steady flow model for flow separation in the representation proposed by Ishizaka and Flanagan (1972), because flow separation would always occur at the sharp edges. Additionally,

- pressure recovery and nonlinearity of the spring are neglected
- Because the contraction formed by the entrance of the glottis is gradual, the vena contracta effect, which is expected to occur in the case of an abrupt contraction with sharp edges, is neglected.

The equations of motion read:

$$m_1\ddot{x}_1 + r_1\dot{x}_1 + k_1x_1 + k_c(x_1 - x_2) = F_1 \tag{2.14}$$

$$m_2\ddot{x}_2 + r_2\dot{x}_2 + k_2x_2 + k_c(x_2 - x_1) = 0 \tag{2.15}$$

where $F_1 \propto \int_{\text{inlet}}^{\text{separation}} P(x)\,\mathrm{d}\vec{S}$ is the novel implementation of the pressure force according to the theoretical prediction of flow separation and the more smooth geometry. No force is assumed to be exerted on the second mass: this way, the second mass "behaves as a child which is held by the hand by a running parent" (Lous et al., 1998), i.e. it is always out of phase with respect of the first mass. This results into the net energy transfer discussed in Section 2.1.2.

Closure of the vocal cords A modified description of vocal fold closure is proposed: vocal folds gradually collide from the lateral side to the middle part, and in some cases (as breathy voice for instance) the glottis never closes completely (Pelorson et al., 1994). As a result of that, during collision there can be still a certain amount of air which is allowed to flow through the glottis (for details, see Pelorson et al., 1994). An important consequence for the quality of the signals is that the resulting spectrum is much less rich in high harmonics which in the rectangular model (Ishizaka and Flanagan, 1972) yield

the typical "computer accent" in numerical program for speech production. The model, together with a more realistic description of the vocal fold collision, can lead to signals surprisingly close to those observed in real speech by inverse filtering.

Steinecke and Herzel (1995)

The model of Ishizaka and Flanagan (1972) is analyzed in this work with the methods of nonlinear dynamics (see Section 2.2). In particular, the attention is directed to the effect of a tension imbalance between the two laryngeal sides, i.e., different stiffness values of the left and right vocal fold that are characteristic for laryngeal paralysis. Because this model is not an attempt to reproduce vocal folds vibration or normal voice output, the aerodynamics is simplified by assuming a steady flow-separation point during convergent glottis, and pressure drop due to viscous losses are neglected. The Bernoulli equation reads:

$$P_s = P_1 + \frac{\rho}{2}\left(\frac{U}{a_1}\right)^2 = P_{\text{sup}} + \frac{\rho}{2}\left(\frac{U}{a_{min}}\right)^2 \tag{2.16}$$

where $a_{min} = \min\{a_1, a_2\}$. The pressure on the second mass P_2 is assumed to be zero. By means of eq. 2.16 (and by setting $P_{\text{sup}} = 0$) the pressure on the first mass P_1 reads:

$$P_1 = P_s\left(1 - \theta(a_{min})\left(\frac{a_{min}}{a_1}\right)^2\right)\theta(a_1) \tag{2.17}$$

where θ is the *Heaviside* function defined as $\theta(x) = 1$ if $x > 0$, $\theta(x) = 0$ otherwise. Consequently, as in previous rectangular models (Ishizaka and Flanagan, 1972), the pressure force reads $F_1 = \ell d_1 P_1$. An additional collision force C_i is implemented as in (Ishizaka and Flanagan, 1972), but neglecting the nonlinear term, thus $C_i = -\Theta(-a_i)c_i\frac{a_i}{2l}$. The equations of motion read:

$$m_1\ddot{x}_1 + r_1\dot{x}_1 + k_1x_1 + k_c(x_1 - x_2) = F_1 + C_1 \tag{2.18}$$

$$m_2\ddot{x}_2 + r_2\dot{x}_2 + k_2x_2 + k_c(x_2 - x_1) = C_2 \tag{2.19}$$

Lous et al. (1998)

Lous et al. (1998) implemented a two-mass model derived from Pelorson et al. (1994) with applications to prosthesis design. This leads to several modifications due to, e.g., size reduction (the prosthesis would probably be much smaller than the vocal folds) and implantation location (the prosthesis could not be placed exactly at the original vocal fold position). The model structure is assumed to be symmetrical ($m_1 = m_2$, $k_1 = k_2$), and the parameters value are not only dictated by physiology, but they also have to make up for over-simplifications in the model. For instance, the total vocal fold mass is slightly increased to compensate for the absence of inlet losses as compared

to Ishizaka and Flanagan (1972). The spring constants k_i are determined from analysis of the modes of vibrations of the mechanical system, and the damping coefficients r_i are chosen according to the values reported in Ishizaka and Flanagan (1972). Collision forces are modelled in a similar way to Pelorson et al. (1994) by considering a stepwise increase in the values of k_i and r_i. The equations of motions read as eq. 2.14-2.15 but in addition the aerodynamic force $F_2 \neq 0$. F_i is the component along the m_i displacement of the force generated by considering a Bernoulli effect and a torque balance on each of the mass-less plates. Acoustic feedback from vocal tract and trachea is considered. The model predictions compare well to *in vivo* experimental data from literature concerning acoustic feedback.

The acoustic properties of the model proposed by Lous et al. (1998) are examined in Sciamarella and D'Alessandro (2004), where an algorithm is developed in order to study the sensitivity of acoustic parameters to the variation of the model control parameters in order to describe the actions that the modelled glottis employs to produce voiced sounds of different characteristics.

A simplified version of the model of Lous et al. (1998) will be presented in Chapter 3 to model the syrinx of the ring dove (*Streptopelia risoria*).

Body-cover model

The so-called body-cover model (Titze, 1988; Story and Titze, 1995) considers separately the body of the folds (muscular and deep ligament layer) and the cover (outer tissue layer of the vocal folds).

In Titze (1988), a framework of basic principles is introduced along the lines of Section 2.1.2. The model is presented within a small-amplitude analytical treatment (the more general large-amplitude theory, which involves limit cycles and nonlinear effects due to collision of the folds is not discussed) to study oscillation threshold pressure which is reduced by reducing the mucosal wave velocity, by bringing the vocal folds closer together and by reducing the convergence angle in the glottis. The body is considered to be stationary, and the cover is assumed to propagate a surface wave along the glottis. This reduces, with respect to previous models as Ishizaka and Flanagan (1972), the number of viscoelastic parameters, and produces a delayed reacting pressure on the bottom of the folds due to a retarded motion at the top, which seems to be the primary mechanism for a velocity-dependent driving force (Titze, 1988), as discussed in Section 2.1.2. The same retarded motion is achieved in two-mass models by a flow controlled mode which keeps the masses out of phase (Ishizaka and Matsudaira, 1972).

In Story and Titze (1995), the vertical mucosa wave propagation is represented, as in classical two-mass models (Ishizaka and Flanagan, 1972) by a linear spring. Additionally, the two masses are coupled laterally to a third mass (the "body mass") by nonlinear

springs and viscous damping elements. The body mass (representing muscle tissue) is further coupled laterally to a rigid wall (representing thyroid cartilage) The model is proposed as a research tool to study both normal and pathological vocal-fold vibrations. Simulations show reasonably similarity to observed vocal-fold motion, measured vertical phase difference, mucosal wave velocity and experimentally obtained intraglottal pressure. Equations of motions include a further variable x_b representing the body mass.

2.1.4 Acoustics of speech

We have reviewed some significant two-mass models of the human vocal folds. In all models, above a threshold value, the pressure causes a net energy transfer from the airflow to the folds that induces self-sustained oscillations of the masses. The opening and closing of the glottal aperture creates pressure waves, i.e., sound. We review here some basic concept of acoustics and vocal tract filtering, introducing to the processes which take place when the sound produced in the glottis travels the vocal tract before the radiation from the mouth.

Sound pressure disturbance is based on local oscillations in an elastic medium, and involves a minimum of two independent variables, space and time, which can be represented by a general function $p = f(x,t)$, called *waveform* or *wave*. Acoustic waves are mathematically the solutions of the acoustic-pressure wave equation $\frac{\partial^2 P}{\partial t^2} = c^2 \frac{\partial^2 P}{\partial x^2}$, where c is the speed of sound. Two highly symmetric and throughly discussed solutions are the *plane wave* $p(x,t) = Ae^{i(\frac{2\pi}{\lambda}x - \omega t)}$ and the *spherical wave* $p(x,t) = \frac{A}{r}e^{i(\frac{2\pi}{\lambda}r - \omega t)}$, where $\omega = 2\pi\mathscr{F}$ is the angular frequency, $\lambda = 2\pi\frac{c}{\omega}$ is the *wavelength*, r is the radius of the vector x and A is called *amplitude* (Mindlin and Laje, 2005).

A waveform is not always an efficient way to display data because there can be much redundancy (Titze, 1994). An alternate approach is to use a spectrum (or spectrum of frequencies): given any sound, we can ask about the number of component frequencies present in the sound and the strength of each component.

Fourier transform

Mathematically, given a function $f(t)$[3] with period T, f can be written as:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\omega_n t) + \sum_{n=1}^{\infty} b_n \sin(\omega_n t), \tag{2.20}$$

[3] The applicability of the Fourier transform depends on mathematical hypothesis which are not discussed in this text.

where

$$a_k = \frac{2}{T}\int_T f(t)\cos\omega_k t\,\mathrm{d}t,$$
$$b_k = \frac{2}{T}\int_T f(t)\sin\omega_k t\,\mathrm{d}t.$$

are termed *Fourier coefficients*. Equation 2.20 is called Fourier series of the function f and represents f in terms of a sum of functions at discrete frequencies $\omega_n = \frac{2\pi n}{T}$.

Alternatively, eq. 2.20 can also be written in a complex notation with Fourier coefficients $\varphi_n = \frac{a_n - ib_n}{2}$:

$$f(t) = \sum_{n=-\infty}^{+\infty} \varphi_n e^{i\omega_n t}, \tag{2.21}$$
$$\varphi_n = \frac{1}{T}\int_T f(t)e^{-i\omega_n t}\,\mathrm{d}t. \tag{2.22}$$

For the general case of a function not necessarily periodic, the equivalent of eq. 2.21 (obtained for $T \to \infty$ and therefore by replacing the discrete set ω_n with the continuous set ω) reads:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \varphi(\omega)e^{i\omega t}\,\mathrm{d}\omega, \tag{2.23}$$
$$\varphi(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t}\,\mathrm{d}t, \tag{2.24}$$

where the coefficients φ_n become a function $\varphi(\omega)$ called *Fourier transform* of the function f. We call *amplitude spectrum* the function:

$$H(\omega) = |\varphi(\omega)| = \sqrt{\varphi(\omega)\varphi^*(\omega)} = \sqrt{\mathrm{Re}^2(\varphi(\omega)) + \mathrm{Im}^2(\varphi^*(\omega))} \tag{2.25}$$

Therefore, for a periodic signal the amplitude spectrum will be an infinite discrete collection of values $(H_1, H_2, \ldots)$ where

$$H_j = H(\omega_j) = |\varphi(\omega_j)| = \sqrt{\varphi(\omega)\varphi^*(\omega)} = \frac{1}{2}\sqrt{a_j^2 + b_j^2} \tag{2.26}$$

represent the amplitudes of the frequency components $(\omega_1, \omega_2, \ldots)$ given in eq. 2.20. For most voice signals, ω_1 is said to be *fundamental frequency* of the signal f, whereas ω_j, $j > 1$ are said to be *harmonics* or *overtones*[4].

Summarizing, a periodic wave contains a discrete quantity of frequencies components $(\omega_1, \omega_2 = 2\omega_1, \ldots, \omega_n = n\omega_1, \ldots)$ with relative amplitudes H_i defined above. This

[4]The frequency components $(\omega_1, \omega_2, \ldots)$ are usually denoted as $(f_0, f_1, \ldots)$ or $(F_0, F_1, \ldots)$. Except for this section, throughout this thesis we will use the notation $(\mathscr{F}_0, \mathscr{F}_1, \ldots)$ and relative amplitudes $(H_0, H_1, \ldots)$.

is the general case of the pressure disturbance radiated from the mouth during speech or singing in normal phonetic conditions. For a non periodic wave no fundamental frequency ω_1 is present (and consequently no harmonics $n\omega_1$), but a continuous function where the relative amplitudes are not in relation with each other.

The plot of the amplitudes versus frequencies is called *spectrum*. For a spectrum of a periodic wave, the *spectral slope* is defined as a measure of how the amplitudes of successive frequencies components ω_i decrease with increasing harmonic number (Titze, 1994). For instance, a sound perceived as "brassy" would have a relative spectrum with a smaller spectral slope than a "fluty" sound (Titze, 1994). Moreover, spectral slope relates to the widely discussed topic on how many birds can produce almost pure tone sounds with similar biomechanical principles as the vocal folds (see Chapter 3 and Section 2.3). A pure tone is defined as a wave with just one frequency component (the fundamental), e.g., a sine function.

The *power spectrum* is defined as the function $H^2(\omega) = |\varphi(\omega)|^2$, by analogy to the case when the ordinate represents power (i.e., energy per unit time): consider a signal originating in the reception of waves (e.g., sound) by an antenna. If the detector is linear (and does not introduce distortion in the frequency band considered) then the quantity it measures is proportional to the variation of the electric field or of the pressure in the vicinity of the antenna (Berge et al., 1986). The theory of waves shows that the acoustic power $\mathscr{P}$ (the amount of energy produced and radiated into the air per second) carried by a wave is proportional to the square of its amplitude averaged over time. We can replace the average over time by an average over frequency by means of the *Parseval-Plancherel* equation

$$\int_{-\infty}^{+\infty} |f(t)|^2 \, \mathrm{d}t = \int_{-\infty}^{+\infty} |\varphi(\omega)|^2 \, \mathrm{d}\omega \tag{2.27}$$

which states that the total energy contained in a waveform $f(t)$ summed across all of time t is equal to the total energy of the waveform's Fourier Transform $\varphi(\omega)$ summed across all of its frequency components ω.

According to these results, in Chapter 3 we will use a quantity called *Harmonic Ratio* (HR) to compare the energy concentrated at successive harmonic frequencies of an acoustic signal. In other words, the HR is an estimation of the spectral slope of the signal: assumed a reference unit (the energy concentrated at the fundamental frequency, H_1^2), the HR value (calculated until the second harmonic only) will be:

$$\mathrm{HR} = 10 \log \frac{\mathrm{H}_2^2}{\mathrm{H}_1^2} \tag{2.28}$$

This quantity is related to the *sound pressure level* (SPL) defined as

$$\mathrm{SPL} = 20 \log \frac{P}{P_{\mathrm{ref}}} \tag{2.29}$$

(compare eq. 2.28 and the above mentioned relation [energy] $\propto$ [pressure]^2) where P_0 is a standard reference pressure of $2 \cdot 10^{-5}$ Pa (which is usually considered the threshold of human hearing). However, because different reference levels P_{ref} can be used, it is necessary to know the reference level before comparing measurements of SPL (Titze, 1994).

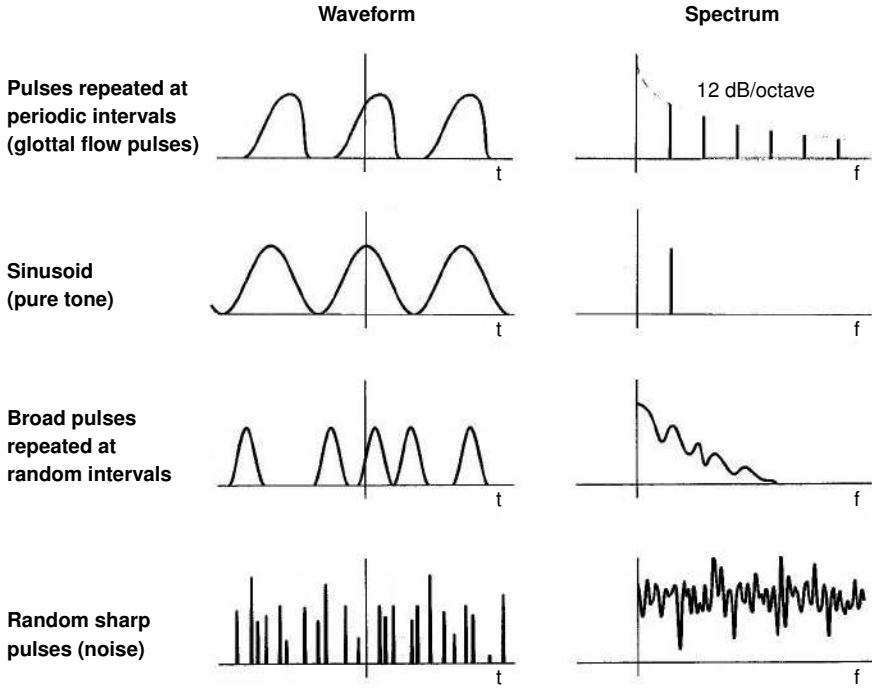


Figure 2.4: Waveforms and relative power spectrum of some typical signals. Power spectra are usually represented on a logarithmic scale on the y axis (modified after Titze, 1994).

Spectrogram The power spectrum of a signal gives information about the signal frequencies content and the energy at each frequency component (Fig. 2.4). However, if the signal frequencies content varies in time, it is useful to represent it as a spectrogram. Spectrograms are based on many subsequent short-time power spectra of overlapping signal segments. For each of these segments, the power spectrum is computed at the relative time (x axis), the spectral amplitudes are encoded as a grey or colored scale, and the frequencies are plotted on the y axis (see, e.g., Table 2.3 on p. 37).

Travelling the vocal tract

Beside the differences often emphasized between sound production in humans and birds, it is important to recognize also the similarities: in both the cases the acoustic output of a vibrating valve (i.e., the power spectrum of the sound pressure radiated by the valve) is modified by vocal tract resonances (Fletcher and Tarnopolsky, 1998). A *resonance* is the constructive interference of waves experiencing multiple reflections in

tube	**formants** ($n = 1, \ldots, \infty$)
open end	$F_n = (2n-1)\frac{c}{4\ell}$
closed end	$F_n = n\frac{c}{2\ell}$

Table 2.1: Formant frequencies for an ideal tube open at the end (open end) and closed at both ends (closed end). ℓ: length of the tube, c: sound velocity. Although there are infinite formants, only a few formant frequencies are usually of interest. For instance, vowels in human speech are perceived and classified on the basis of the first two formant frequencies of the vocal tract (Peterson and Barney, 1952; Bogert and Peterson, 1957; Kent, 1978, 1979).

a tube (Titze, 1994). The human vocal tract, like the tube of a wind or brass instrument, resonates at certain special frequencies (called *formants*). The resonant frequencies depend on the shape of the vocal tract, e.g., the configuration of tongue, jaw and lips in the case of humans, and tongue and beak in the case of birds (Fletcher and Tarnopolsky, 1998). In humans, formants determine many speech sounds (phonemes) from which syllables and words are made. For an idealized tube, the formant frequencies are given in Table 2.1.

The acoustic properties of the human vocal tract have traditionally been described on the basis of a *source-filter theory* (Fant, 1960). In this theory, the sound source is the time-varying glottal airflow and the filter is the vocal tract. Whereas the glottis produces a sound of many frequencies, the vocal tract selects (filters) a subset of these frequencies for radiation from the mouth (Titze, 1994). The source filter theory assumes that the labial dynamics are independent of whatever happens in the filter. It has recently been shown, however, that this assumption is generally not valid, and only under certain conditions is at best an appropriate simplification. The linear theory is applicable to male speech, less to female and child speech, and much less to singing (Titze, 2007). The source-filter hypothesis, i.e., a passive linear filter of the vocal tract, is based on the hypothesis that the glottal output impedance is much larger than the vocal tract input impedance[5]. This hypothesis holds under normal speech conditions: the mean glottal area is much smaller than the input cross section of the vocal tract, and the fundamental frequency of the glottis is below the first formant of the tract (Laje et al., 2001). Laje et al. (2001) assumed that P_{sup} is the result of adding perturbations $P(t)$ induced by labial

[5]The concept of *wave impedance* (or *specific acoustic impedance*) Z can be used to discuss reflections of sound from interfaces among various media (e.g., human voice: Titze, 1994; bird songs: Mindlin and Laje, 2005; musical instruments: Fletcher and Rossing, 2005). Z quantifies the amount of pressure P that is produced for a given speed of the air particles v:

$$Z = \frac{P}{v}$$

and is measured in Pa sec m^{-1}.

vibration and the perturbations arriving back after being reflected by the tract: $P_{\text{sup}} = P(t) + P_{back}(t-\tau)$, where τ is the time it takes for a sound wave to travel the length of the tract and $P(t) = I\dot{U}$ (Rothenberg, 1981; Titze, 1994). For large enough values of the inertance I (which in turn determines the amplitude of P) the pressure fluctuations that are established at the base of the trachea can be important and affect labial dynamics. Beside this (nonlinear) interaction of glottal airflow with acoustic vocal tract pressures, a more dramatic level of interaction occurs in high fundamental frequencies productions for which the dominant (first four) harmonics cross the formants (Titze, 2007): a variety of new source frequencies can then be produced, including subharmonics, frequency jumps, and deterministic chaos, which can be explained with the theory of nonlinear dynamics.

2.2 Nonlinear dynamics of the vocal folds

The vocal folds, together with glottal airflow, constitute a nonlinear oscillating system. This system is amenable to nonlinear dynamics analysis, which can provide an entirely different perspective on the underlying basis of phonatory function (Titze et al., 1993).

In recent years it has been suggested that phenomena reported for decades in the phonetic literature such as “octave jumps”, “biphonation” (two independent frequencies) and *noise* concentration, could be interpreted as period doubling bifurcations, tori and chaos. Subharmonics and irregularities can appear occasionally in normal phonation such as newborn cries (Mende et al., 1990). On the other hand, vocal instabilities due to bifurcations are often symptomatic in voice pathology (Herzel and Wendler, 1991; Titze et al., 1993; Herzel, 1993; Herzel et al., 1994), and nonlinear phenomena such as period doubling, biphonation and irregular aperiodic phonation are exploited in contemporary vocal music (Neubauer et al., 2004). In this framework, several models have been developed in recent years to study vocal fold paralysis (Mergell et al., 2000), instabilities at high pressures (Jiang et al., 2001), unilateral laryngeal paralysis by means of a tension imbalance (Steinecke and Herzel, 1995), aperiodic vibrations of a vocal fold model with a polyp (Zhang and Jiang, 2003) or instabilities such as abrupt jumps between chest-like and falsetto-like vibrations exhibiting hysteresis, aphonic episodes, subharmonics, and chaos near the register transitions (Tokuda et al., 2007).

While on the one hand human vocal production demonstrates that certain irregular phenomena seen in human pathological voices and baby crying result from nonlinearities in the vocal production system, on the other hand equivalent phenomena are quite

common in nonhuman primate vocal repertoires (Wilden et al., 1998; Fitch et al., 2002; Tokuda et al., 2002; Neubauer, 2004). In particular, bifurcations and chaos are ubiquitous aspects of the normal adult repertoire in many, if not most, primate species (Fitch et al., 2002). An analysis of vocalizations of macaque screams, piglet screams, and dog barks has shown that nonlinear analysis is primarily useful in animal vocalizations (including subharmonics and biphonation) or low-dimensional chaos (Tokuda et al., 2002).

In birds literature, as we shall see later, it has been recently observed that biomechanical principles similar to human phonation underlie bird song production. Recently (Fee et al., 1998; Fee, 2002), nonlinear dynamics has been studied in a modified two-mass model to study frequency-mode locking and register transitions in the vocalizations of the zebra finch (*Taeniopygia guttata*). Gardner et al. (2001), Laje et al. (2002), Laje and Mindlin (2002), Mindlin et al. (2003) and Laje and Mindlin (2005) developed several models based on the mucosa-wave model of Titze (1988) to study nonlinear effects caused by control parameters in birdsongs.

In this section, following standard textbooks (Guckenheimer and Holmes, 1983; Berge et al., 1986; Jackson, 1989), the basic principles of (nonlinear) deterministic dynamical systems will be reviewed.

2.2.1 Introduction to nonlinear dynamics

Formally speaking, we define *dynamical system* a system whose behavior can be described by an *evolutional operator* $\Phi^t : X \to X$ defined on a space X (*phase space*) $\forall t \in \mathscr{T}$ (in the remainder of this section, we will assume $X = \mathbb{R}^{\mathrm{n}}$, $\mathrm{n} \geq 1$). If the space $\mathscr{T}$, or time, is $\mathbb{R}$, the system is said to be continuous dynamical system, if $\mathscr{T} = \mathbb{Z}$, the system is said to be discrete dynamical system. Two conditions must hold for Φ for any initial condition $x_0 \in X$ and any $t, s \in \mathscr{T}$: the above mentioned deterministic property, expressed by the relation $\Phi^{t+s}(x_0) = \Phi^t(\Phi^s(x_0))$, and the property of "no evolution at time zero": $\Phi^0(x_0) = x_0$.

In a continuous dynamical system the evolution in time of the points of X is typically given as a vector field or *flow*:

$$\dot{x} = f(x) \tag{2.30}$$

where $f : X \to X$. In *nonlinear* dynamical systems eq. 2.30 is nonlinear, in other words the variables x_i, $\dot{x}_i$ in eq. 2.30 form product terms or contain higher order terms. Contrarily to linear differential equations, there is no general way to obtain families of solutions of nonlinear equations (except when they exhibit symmetries): linear equations frequently appear as approximations to nonlinear equations, and these approximations are only valid under restricted conditions. Any (deterministic) nonlinear dynamical system is potentially chaotic, i.e., even though "determined" entirely by internal properties and initial conditions, can exhibit, over the long range, unpredictable behavior (*chaos*).

For the case at hand, let us consider a two-mass model based on the classical mass-spring oscillator (e.g., eq. 2.18 and 2.19 on p. 17, denoting for simplicity $f_1 = F_1 + C_1$ and $f_2 = C_2$). The dynamics of the two moving points are represented by two second-order differential equations which can be reduced to a system of four first-order equations, thus leading to the vector field:

$$\dot{x} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} x_2 \\ \frac{1}{m_1}\left(f_1 - r_1 x_2 - k_1 x_1 - k_c(x_1 - x_3)\right) \\ x_4 \\ \frac{1}{m_2}\left(f_2 - r_2 x_4 - k_2 x_3 - k_c(x_3 - x_1)\right) \end{pmatrix} = f(x) \tag{2.31}$$

(compare eq. 2.30) defined on the phase space $X = \mathbb{R}^4$. A generic point

$$x = \vec{x} = (x_1(t), x_2(t), x_3(t), x_4(t)) \in X$$

will represent, at instant t, the position and the velocity of the first mass, x_1 and x_2 respectively, and the position-velocity of the second mass (x_3 and x_4 respectively). Eq. 2.31 can be written in the general compact form

$$\dot{x} = f(x, \lambda) = f(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) \tag{2.32}$$

where $\lambda \in \mathbb{R}^m$ is a vector of m parameters (e.g., k_i, r_i, k_c in eq. 2.31) and $x \in \mathbb{R}^4$.

Nonlinearity is found in vocal fold mechanics in *at least* two ways (see eq. 2.31):

- highly nonlinear interaction of the airflow and glottal area (van den Berg et al., 1957).
- some quadratic or cubic relation, such as the cubic term in eq. 2.10 to model tissue elastic deformation and collision force (Ishizaka and Flanagan, 1972), or the nonlinear dissipation term $cx^2\dot{x}$ in the flapping model (Laje et al., 2002; Laje and Mindlin, 2002).

Special solutions of a vector field (if they exist) are:

① *Equilibria* or *fixed (or stationary) points*: points $x^* \in X : f(x^*) = 0$

② *Periodic orbit*: the closed curve $\Phi^t(x) : 0 \leq t < T$, where T, called *period* is the smallest number such that $\Phi^T(x) = x$

In the time evolution of the dynamical system, starting from initial state $x_0 \in X$, the phase flow f generates a set of states $\{\Phi^t(x_0) \in X, 0 < t < \tau\}$ ($\tau \in \mathbb{R}$) called *trajectory* in phase space, depending on the associated initial state. Trajectories can never intersect in the phase space, which is an unavoidable consequence of the deterministic nature of the description (Berge et al., 1986). Trajectories can asymptotically converge into sets

of phase space points called *attractors*, a bounded attractive set of phase space points. Given an attractor $\mathscr{A}$, the set of points of the phase space X which approaches asymptotically to $\mathscr{A}$ is termed *basin of attractor* of $\mathscr{A}$. Points already within the attractor set are mapped to points within the attractor, i.e. an attractor is invariant under the phase flow f. Trajectories starting from a general $x_0 \notin \mathscr{A}$ in the basin of attractor of $\mathscr{A}$ will approach $\mathscr{A}$ after a period of time called *transient*. As shown in Fig. 2.5, attractor sets can be subdivided into four different classes (Jackson, 1989):

- **Steady states**: steady states are represented by fixed points in phase space. For instance, the trajectories of a damped harmonic oscillator approach asymptotically the stable stationary rest state.

- **Limit cycles**: limit cycles are isolated closed curves in phase space (as in, for example, the asymptotic behavior of a driven, damped harmonic oscillator).

- **Torus**: Trajectories associated with quasiperiodic functions (r independent frequencies) lie on a r-dimensional torus, T^r. E.g., if $r = 2$, a torus is a two-dimensional object in phase space, where the trajectory of two coupled oscillators asymptotically moves on the surface, never meets itself again and asymptotically fills the whole surface.

- **Chaotic attractors**: a chaotic attractor is typically a fractal set of points in phase space. Stressing the fractal nature of this set, it is termed strange attractor. Stressing sensitive dependence of single trajectories on initial conditions and exponential divergence of initially close trajectories, the term "chaotic attractor" is used (Strogatz, 1994). A single trajectory on a strange attractor is aperiodic, irregular, and deterministic, but not random or stochastic.

Given a complete description and a precise measurement of dynamical variables and parameters of, e.g., the phonatory system, attractor and bifurcation analysis could reveal all possible different dynamical behaviors of the system. Unfortunately, in real world not all variables are directly measurable and values of system parameters might be inaccurate. Nevertheless, from a signal $x(t)$, a vector can be reconstructed in *pseudo-phase space* by "delay-coordinates" $x(\tau + t)$ (Takens, 1981; Froehling et al., 1981; Eckmann and Ruelle, 1985; Story, 1991; Titze et al., 1993; Guckenheimer, 1999).

For constant control parameters $\lambda_i \in \lambda$ and given initial conditions, the long time behavior of a nonlinear system is uniquely determined by the attractor types associated to the attractor basins. Lyapunov exponents, Poincaré sections and Fourier analysis, as discussed in Section 2.2.4, are often enlightening to classify and recognize attractors. For varying components of the parameter vector λ, the local topology of attractors in phase space can change qualitatively. At critical parameter values, bifurcations between

Attractor type	Sketch of the attractor	Poincaré map	Dimension	Lyapunov exponents
Stationary point			0	$\lambda_1 < 0$
Limit cycle			1	$\lambda_1 = 0$ $\lambda_2, \ldots, \lambda_n < 0$
k-torus	 (2-torus)		k	$\lambda_1, \ldots, \lambda_k = 0$ $\lambda_{k+1}, \ldots, \lambda_n < 0$
Chaotic attractor			> 2 (typically $\notin \mathbf{Z}$)	$\lambda_1 > 0$

Figure 2.5: Various attractors and their characteristics (modified after Titze et al., 1993).

different attractor types are induced that qualitatively change the asymptotic behavior of trajectories. Stability of attractors (Floquet theory) and bifurcations are briefly discussed in the next section.

2.2.2 Stability of fixed points, Floquet theory and bifurcations

The stability of a fixed points is reduced to the study of the stability of the linearized system in a neighborhood of x_0, thus calculating the eigenvalues of the *Jacobian matrix* $\left[\partial f_i(x)/\partial x_j\right]_{x=x_0}$ (Guckenheimer and Holmes, 1983) and leading to the possible situations depicted in Table 2.2: an attractor (repellor) is a stable (unstable) fixed point, whereas a saddle node is stable in a subspace $U \subset X$ and unstable in $X \backslash U$.

Stability of a periodic orbit can be described by means of Floquet theory of periodic solutions. Given a periodic solution $x(t) = x(t+T)$ on a periodic orbit γ, linear stability analysis (Floquet theorem) assures that there exists a constant matrix R for which a small initial displacement away from the solution, $x + \delta x$, will be of the form $x + e^{TR}\delta x$ after a period T (see Guckenheimer and Holmes, 1983; Berge et al., 1986, for details). The eigenvalues $\rho_1, \ldots, \rho_n$ of e^{TR} are called *Floquet - or characteristic - multipliers*. The periodic orbit is stable if each component of $e^{mTR}\delta x$ on the hyperplane transversal to the

eigenvalues real part (sorted in descending order from left to right)	**fixed point stability**
$(+,\ldots)$	unstable:
	$(+,\ldots,+)$ *repellor*
	$(+,\ldots,+,-,\ldots,-)$ *saddle point*
$(-,\ldots,-)$	stable (*attractor*)
$(0,\ldots,-)$	stability can not be determined by linearization

Table 2.2: Stability types for a fixed point x_0 of a n-dimensional dynamical system according to the sign of the eigenvalues of the Jacobian matrix evaluated in x_0.

orbit direction decreases in time ($m \to \infty$), i.e., if $|\rho_2| < 1, \ldots, |\rho_n| < 1$ (the matrix e^{TR} has always an eigenvalue equal to one which corresponds to a displacement along the trajectory. We suppose this eigenvalue to be ρ_1). The Floquet multipliers $(\rho_2, \ldots, \rho_n)$ represent the eigenvalues of the linearized Poincaré map (see Section 2.2.4). The eigenvalues λ_i of the constant matrix R (not to be confused with λ_i in eq. 2.32) are called *Floquet* (or *characteristic*) *exponents*[6].

Floquet theory needs a representation of e^{TR}, which is generally difficult and requires perturbation methods or the use of special functions. The characterization of attractors is often obtained via Poincaré sections, Lyapunov exponents or Fourier analysis, as discussed later. However, the study of transitions through the loss of linear stability is a convenient way to classify a number of processes through which chaotic behaviour will occur. Among them, *period doubling bifurcation* (a Floquet multiplier crosses the unit circle at -1), or *secondary Hopf bifurcation* (two complex conjugate Floquet multiplier cross the unit circle), as discussed in the next section.

2.2.3 Bifurcations

One central goal of model analysis has been in the past the determination of the critical subglottal pressure P_{th} for the start of self-sustained oscillations in vocal folds (Mergell et al., 1998): the change in the dynamics behavior is based on a *Hopf bifurcation*, i.e., at a critical pressure value a stable equilibrium of the system becomes unstable and the dynamics undergo a change from a stable fixed point to a limit cycle attractor. Generalizing: whenever the solution to an equation (or system of equations) changes quali-

[6]Note that, $\forall j = 1, \ldots, n$, $|\rho_j| = |e^{\lambda_j T}|$, where $\lambda_j = a + ib \in \mathbb{C}$. From the equivalence

$$|\rho_j| = |e^{(a+ib)T}| = |e^{aT}(\cos bT + i \sin bT)| = |e^{aT}| = e^{2aT} = e^{2T\mathrm{Re}(\lambda_j)}$$

we have: $|\rho_j| < 1 \Leftrightarrow \mathrm{Re}(\lambda_j) < 0$.

tatively at a fixed value of a parameter $\lambda_i \in \lambda$, this will be called a *bifurcation* (Berge et al., 1986). The plot of any characteristic property of the solutions of eq. 2.32 versus a parameter λ_i is termed *bifurcation diagram*.

Hopf bifurcations are called *supercritical* or *subcritical* if the dynamics undergo a change to a stable or unstable limit cycle, respectively. Subcritical Hopf bifurcation is characterized by a coexistence of two stable states (an equilibrium position and a stable limit cycle), leading in vocal folds vibrations to two distinct thresholds: an onset for increasing P_s and an offset (lower than the onset) for decreasing P_s. At both thresholds, the onset and the offset of vocal fold oscillations happen abruptly, and in the intermediate region, jumps from one dynamic regime into the other can be observed (*hysteresis*). Hysteresis has also been found in excised larynx experiments (Berry et al., 1996). Mergell et al. (1998) have shown theoretically how the envelope function $r(t) \simeq a_1 r + a_2 r^3 + \mathcal{O}(r^5)$ describing the series of oscillations maxima and minima of the masses can lead to super- and subcritical bifurcation by adding high order term and incorporating a positive value of the parameter a_2. Lucero (1993, 1999) has studied the hysteresis phenomenon and calculated the onset pressure for the mucosal wave model of Titze (1988) and described equilibria and threshold pressure of the two-mass model of Ishizaka and Flanagan (1972).

For phonatory systems other important bifurcation types are: secondary Hopf bifurcation, period doubling bifurcation, and abrupt onset of deterministic chaos. Secondary Hopf bifurcations are due to an emergence of a second oscillation with an independent (incommensurable) frequency, changing the behavior from a limit cycle oscillation to a toroidal oscillation. At a period doubling bifurcation, an emerging second frequency at half the frequency of a limit cycle oscillation changes a limit cycle attractor to a folded limit cycle attractor. Cascades of period doubling bifurcations or secondary Hopf bifurcations can be precursors to the onset of deterministic chaos (for further routes to chaos see, e.g., Berge et al., 1986).

2.2.4 Lyapunov exponents, Poincaré sections and Fourier analysis

We review here some methods which are often enlightening to classify and recognize attractors, in addition to the phase portraits: Poincaré sections, Lyapunov exponents and Fourier analysis.

Lyapunov exponents

Lyapunov exponents are quantities which characterize the rate of separation of infinitesimally close trajectories. The rate of separation can be different for different orientations of initial separation vector. Thus, there is a whole spectrum $(\gamma_1, \ldots, \gamma_n)$ of Lyapunov exponents, where n is the phase space dimension. It is common to just refer to the largest

one (maximum Lyapunov exponent γ_1, after sorting the spectrum in descending order), because it determines the predictability of a dynamical system. More precisely, γ_1 indicates that sufficiently small deviations from a trajectory grow, in the mean, according to $\exp(\gamma_1 t)$ (Titze et al., 1993). Thus the system is chaotic if $\gamma_1 > 0$. Otherwise trajectories can approach asymptotically a torus, a limit cycle or a stable fixed point (Fig. 2.5) according to the sign of $\gamma_1, \ldots, \gamma_n$ (see Eckmann and Ruelle, 1985, for details). An algorithm for calculating γ_1 from a time series was developed by Wolf et al. (1985) and used in a modified version in Herzel and Wendler (1991) and Titze et al. (1993).

Poincaré sections

Poincaré sections are another useful tool to characterize attractors. Given a trajectory on $\tau \in X$ it can be fruitful to observe the points of intersection of τ with an hyperplane Σ of X (i.e., $\dim(\Sigma) = \dim(X) - 1$). This set of points comprises the *Poincaré section*, that is, a graph in $n-1$ dimensions, where $n = \dim(X)$. Σ can be any hyperplane, but an appropriate choice yields sections that are more easily analyzed. If $x_0, x_1, x_2, \ldots$ are the points of intersection of τ and Σ, the transformation leading from one point to the next is a mapping of Σ into itself:

$$\begin{aligned} T: \quad & \Sigma \to \Sigma, \\ & x_{k+1} = T(x_k) = \ldots = T^{k+1}(x_0) \end{aligned} \tag{2.33}$$

called *Poincaré map*.

Poincaré section and map have the same kind of topological properties as the flow from which they arise (Berge et al., 1986) and simplify the study of continuous flows consistently. First, Poincaré section replaces the continuous-time evolution of eq. 2.30 with a discrete-time mapping (eq. 2.33), thus replacing differential equations with difference equations, which are substantially easier to solve. Second, the quantity of data to be manipulated is greatly reduced. Third, we pass from a flow in n dimensions to a mapping on a space of dimension $n-1$. Attractors in X have all a characteristic equivalence in the space Σ, as shown in Fig. 2.5.

Fourier analysis

Fourier transforms can also be applied to study a trajectory $\tau \in X$. In (relatively) recent years, the rapid development of computational methods has meant that a signal $x(t)$, a continuous function of time, is very often measured by sampling and discretizing. It is therefore common to deal with sampled measured signals, thus a vector of discrete values $(x_0, \ldots, x_n)$ at times $(0, \Delta t, 2\Delta t, \ldots n\Delta t = t_{max})$. A *discrete Fourier transform* of $\vec{x} = (x_1, \ldots, x_n)$ is the operation creating the corresponding discrete series (compare

Section 2.1.4):

$$\hat{x}_k = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} x_j \exp\left(-\frac{2\pi i k}{n} j\right) \tag{2.34}$$

The aspect of the power spectrum, as discussed in Section 2.1.4, clearly reveals how the signal $\vec{x}$ evolves in time:

- a periodic signal is characterized by a spectrum with peaks at frequencies $\frac{n}{T}$, where $n \in \mathbb{Z}$ and T is the period of the signal[7]
- a quasiperiodic signal with r frequencies $\mathscr{F}_1, \ldots, \mathscr{F}_r$ contains components at all frequencies of the form $|m_1\mathscr{F}_1 + \ldots, m_r\mathscr{F}_r|$ with $m_i \in \mathbb{Z}$[8]
- aperiodic signals are characterized by a continuous Fourier spectrum.

However, the aspect of a continuous spectrum is not distinguishable from a power spectrum generated from an aperiodic signal with a really high number of frequencies. Fourier transform is therefore limited with respect to methods of analysis such as Poincaré sections or Lyapunov exponents. In Chapter 3, in order to confirm the appearance of period doublings and deterministic chaos in the trapezoidal model, we will use both Fourier spectra and modified Poincaré sections.

2.3 Sound production in birds

Along the lines of the section dedicated to humans, we review now basic principles of avian sound production.

Birds share with other terrestrial vertebrates the ability to utilize airflow within the respiratory system as a source of energy (Brackenbury, 1982). The larynx is also present in birds, located at the *rostral*[9] end of the trachea. However, in contrast with mammals, the larynx in birds does not house vocal folds for sound production. The avian vocal

[7] In fact, this situation is rather academic: in practise the period T of the signal is often unknown, the ratio $t_{max}/T \notin \mathbb{Z}$, resulting in a spectrum with secondary peaks centered at each harmonic component, called *sidelobes* (see Berge et al., 1986, p. 50-53).

[8] In case of $r = 2$, $|m_1\mathscr{F}_1 + m_2\mathscr{F}_2|$ is dense if $\frac{\mathscr{F}_1}{\mathscr{F}_2} \notin \mathbb{Z}$, otherwise not. In this latter case, when $\frac{\mathscr{F}_1}{\mathscr{F}_2} = \frac{n_1}{n_2}$ with $n_1, n_2 \in \mathbb{Z}$, the signal is in fact periodic with period $T = n_1T_1 = n_2T_2$. One says that there is a *frequency locking* of $\mathscr{F}_1$ with $\mathscr{F}_2$. All the peaks of the Fourier spectrum are harmonics of the lowest frequency $\mathscr{F}_0 = \frac{1}{T} = \frac{\mathscr{F}_1}{n_1} = \frac{\mathscr{F}_2}{n_2}$ (Berge et al., 1986, p. 57-58).

[9] Situated toward the beak. Rostral is technical Latin terminology, very common in functional morphology and medicine (cf. e.g., *caudal*: situated toward the tail).

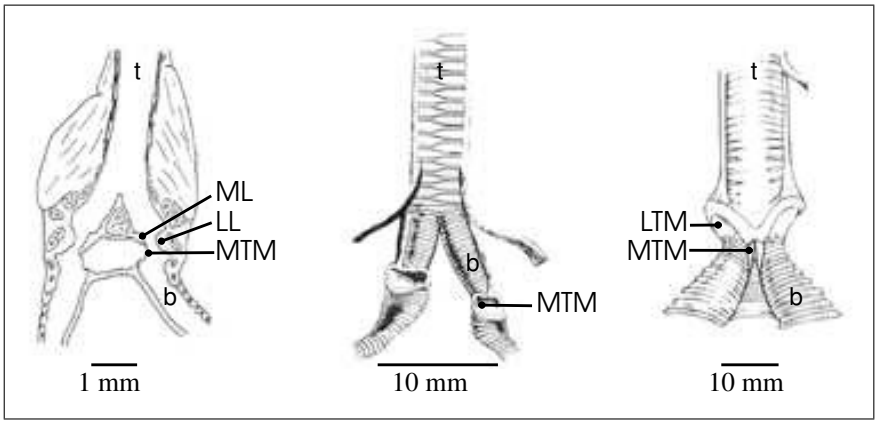


Figure 2.6: Variation in syrinx morphologies (modified after Elemans et al., 2003). From left to right: 1) Brown thrasher (*Toxostoma rufum*), which is a typical Passeriniform syrinx (modified after Suthers et al., 1999 in Elemans et al., 2003). 2) Oilbird (*Steatornis caripensis*, modified after King, 1989 in Elemans et al., 2003) and 3) cormorant (*Phalacrocorax carbo*, modified after King, 1989 in Elemans et al., 2003). ML, Medial Labium; LL, Lateral Labium; LTM, lateral tympaniform membrane; MTM, medial tympaniform membrane; b, bronchus; t, trachea.

organ, unique in the animal kingdom, is called *syrinx* and it is located at the bifurcation of the trachea into the two bronchi.

2.3.1 Anatomy of the avian syrinx

Due to the high metabolic rate required for flying, birds have a high oxygen demand. They meet this by having a respiratory system consisting in relatively small lungs plus nine air sacs that play an important role in respiration. One of this air sacs, the interclavicular air sac (ICAS), completely surrounds the syrinx, which is therefore located within air space.

Enormous morphological diversity is found in the syrinx of different species. Three types of syrinx are recognized (Fig. 2.6): tracheal, bronchial and tracheo-bronchial (for a review of avian syringeal anatomy, see Casey and Gaunt, 1985, and references therein). Tracheo-bronchial syrinx is particularly common in *oscines* (*songbirds*). Songbirds syrinx is a bipartite structure whose two sides are potentially capable of acting either independently or together to produce sound, and in the latter case each side of the syrinx may produce a different sound (Suthers, 1990). For instance, in the northern cardinal (*Cardinalis cardinalis*) and in the brown-headed cowbird (*Molothrus ater*), high-frequency sounds are produced primarily in the right bronchus, and low-frequency sounds are produced primarily in the left bronchus (Suthers et al., 1999). Within the syrinx there are two, distinct types of membranes, the lateral tympaniform membrane (LTM), the medial tympaniform membrane (MTM) and soft tissue masses as the lateral labia (LL) and medial labia (ML). Among the muscles associated with the syrinx, two

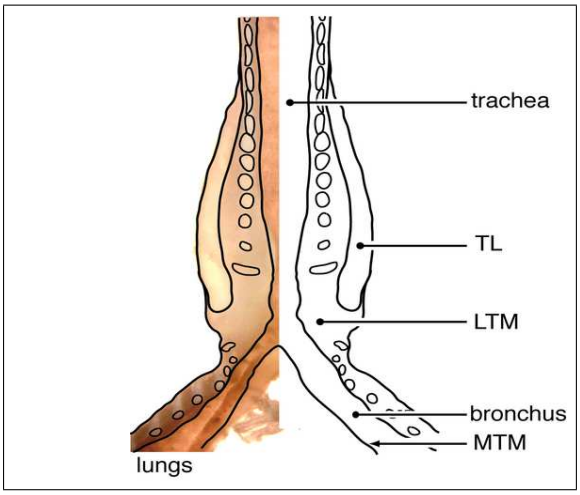


Figure 2.7: Sketch of the syrinx of the ring dove (*Streptopelia risoria*, modified after Zaccarelli et al., 2006).

are extrinsic (*musculus sternotrachealis*, ST, and *musculus tracheolateralis*, TL) arising on the costal process of the sternum (ST) and the cartilages of the syrinx (TL) and inserting on the trachea (Brackenbury, 1982).

Our attention will be pointed in particular on the syrinx of the ring dove (*Streptopelia risoria*), a tracheal syrinx where the vibrating membranes (Fig. 2.7) are located in the tracheal walls immediately cranial to the tracheobronchial junction (Gaunt et al., 1982). Contrarily to songbirds, ring dove syrinx is made of one single (and not two independent) vibrating structure: the LTM is thought to be the primary sound source, while TL and ST play an important role in the phonatory position of the syrinx (Goller and Larsen, 1997b).

2.3.2 Sound production in birds

Mechanism underlying sound production in birds is still poorly understood. In contrast with the situation for the human larynx and vocal tract, for which numerous biomechanical, aerodynamic and acoustic studies exist, there were few attempt to analyze and synthesize the behavior of the avian syrinx and the vocal tract in quantitative terms (Fletcher and Tarnopolsky, 1998). There are obvious reasons for this difference: the great variety of anatomical birds dimensions and vocal utterances, and a strong interest in understanding the functioning of our own body. Moreover, due to its inaccessible location in the birds body, the function of the syrinx was mainly derived from morphological observations and physiological measurements such as flow, pressure and muscle activity (Elemans, 2004) (for reviews, see Nottebohm, 1976; Gaunt and Gaunt, 1985; Goller and Larsen, 2002, and references therein). Despite that, there exists a very large literature on behavioral, acoustic and physiological aspects of bird song (among

them, Greenewalt, 1968; Gaunt and Gaunt, 1985), and some attempt to examine passive acoustics of the avian vocal tract under principles valid fairly generally among different species (Fletcher and Tarnopolsky, 1998).

For decades, the most widely accepted mechanism of sound production in birds (*classical model*) identified the MTM as the vibrating structure. More precisely, sound generation mechanism was attributed to oscillation of an edge-clamped membrane (the MTM) while the LL were thought to provide constriction of the syringeal lumen (Casey and Gaunt, 1985; Gaunt and Gaunt, 1985, and references therein).

However, Goller and Larsen (1997a,b, 2002) changed this dominant idea and demonstrated that in songbirds sound is produced by oscillations or collisions of the labia (Goller and Larsen, 1997a, 2002). After ablation of the MTM, cardinals and zebra finches were both able to phonate and sing nearly normally. For non-songbirds we cannot yet make general statements since the syringeal morphology is more diverse, and only two species have been studied (Elemans, 2004). However, in pigeons the LTMs appears to be the main sound generators, since disabling the MTM with tissue adesive did not prevent phonation nor change the frequency and amplitude structure of vocalizations (Goller and Larsen, 1997b).

These new perspectives of sound generation in songbirds suggest that the physical mechanism of phonation is very similar to that operating in the human larynx (Goller and Larsen, 2002). However, several aspects must be taken into account, e.g., asymmetry in syrinx labial mass and morphological differences between songbird syrinx and human larynx: the ICAS (which ramifies between the two MTMs) and the pessulus (a cartilaginous tissue connecting the medial walls of the bronchi) could induce pressure connections or structural coupling between the two syringeal sources, respectively (Nowicki and Capranica, 1986a,b). Furthermore, the MTM could play a role in affecting the tension of the ML, especially for the generation of high-frequency sounds (Goller and Larsen, 2002). Therefore, it is likely that in songbirds multiple mechanisms and intermediate forms are employed for some sound.

Whistled sound

The acoustic spectra of most birdsongs contain either a fundamental tone with or without harmonics, or broadband noise with all frequencies present (Casey and Gaunt, 1985). Among these registers, the pure tone quality of some birds (fundamental without harmonics, also referred as *whistled song*) has led to the suggestion that birds produce pure-tone signals in a fundamentally different way as humans. Control of harmonics in bird songs is widely debated (Riede et al., 2004; Klatt and Stefanski, 1974; Nowicki, 1987; Nowicki et al., 1992; Beckers et al., 2003a, 2004; Fletcher et al., 2004; Williams et al.,

1989): some species, such as budgerigars (*Melopsittacus undulatus*), have whistle-like songs and not much energy is found in the harmonics (Nowicki et al., 1992), whereas other species, such as the zebra finch (*Taeniopygia guttata*), display strong harmonics during song or calls (see Table 2.3).

Various mechanisms have been hypothesized to be the acoustic source in whistled songs. Arising from theoretical difficulties in explaining how a vibration-based system (as in the human larynx) could generate tonal sound (Casey and Gaunt, 1985; Gaunt and Gaunt, 1985; Fletcher, 1988), alternatives hypotheses were formulated. Among them, the hole-tone whistle model (Nottebohm, 1976; Casey and Gaunt, 1985): a constriction of the syringeal lumen by membranes or labia was thought to form a narrow slot were vortex formation occurs as air is forced through the narrow opening. This is the same principle which underlies the pure tonal content of human whistle (no vocal fold oscillations) versus voice (vocal fold oscillations)[10].

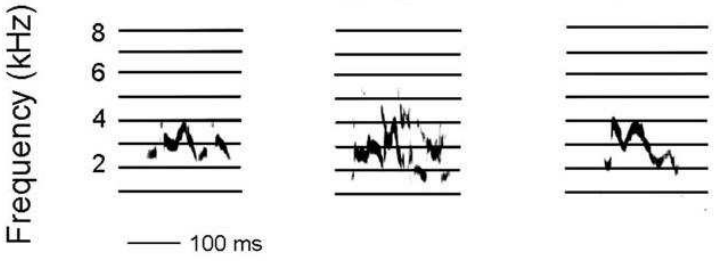
However, experimental evidence by means of songs recorded in heliox atmosphere did not induce substantial shifts of fundamental frequency in any of the investigated species (parrots: Brittan-Powell et al., 1997b; pigeons: Ballintijn and ten Cate, 1998; songbirds: Nowicki, 1987. See Table 2.4). Heliox is a mixture of helium and oxygen (5:1) which speeds up sound velocity. If tonal sounds were generated like tones of a hole-tone whistle, the fundamental frequency should also be shifted upwards in light gas. These findings provide evidence for a vibration-based mechanism, but do not make it possible to discern between the classical model (membrane vibration hypothesis) and alternative models. In this case, the most compelling evidence which undermines the assumption of vibrating membranes came from Goller and Larsen: since then, alternative models of avian sound production were investigated, as discussed later.

In Nowicki (1987), harmonic overtones appeared in vocalizations of nine species of songbirds which were normally pure-tonal. This result could be explained by a mechanism derived from human soprano singing ("soprano model"): an overlap between vocal tract resonance and valve vibration frequencies results in nonlinear feedback that suppresses production of harmonics at the source. However, only slight effects on the spectral content of the vocalizations were observed in parrots and pigeons (Brittan-Powell et al., 1997b; Ballintijn and ten Cate, 1998. See Table 2.4), which suggests

[10] An interesting analogy with the acoustics of musical instruments is discussed in Fletcher (1988): in *voiced sound*, i.e., when there is mechanical motion of part of the vocal system in the course of sound production (classical model), an analogy can be found with the mechanism of sound production in reed-driven woodwind musical instruments (e.g., clarinets, saxophones or oboes). On the other hand *whistled songs*, i.e., sounds produced aerodynamically by means of instabilities of air jets in interaction with apertures, edges and cavities (hole-tone model) may be discussed in analogy with the sound production mechanism in flutes and organ pipes (see Fletcher, 1988, and references therein).

Budgerigars (*Melopsittacus undulatus*)

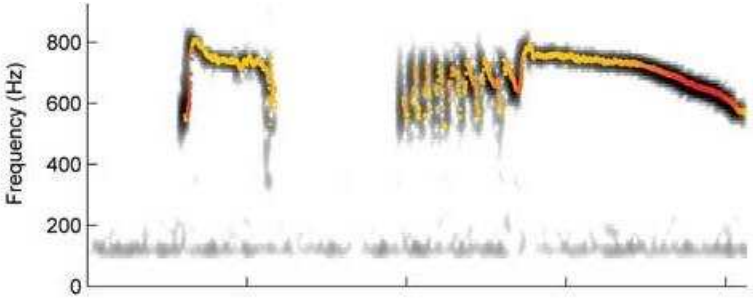
Order: Psittaciformes (parrot or psittacine). Vocal repertoire mainly tonal. (Brittan-Powell et al., 1997b, and references therein).



sources (picture and spectrogram): `http://en.wikipedia.org/wiki/Image:Blue_male_budgie.jpg` and Brittan-Powell et al. (1997a)

Ring dove (*Streptopelia risoria*)

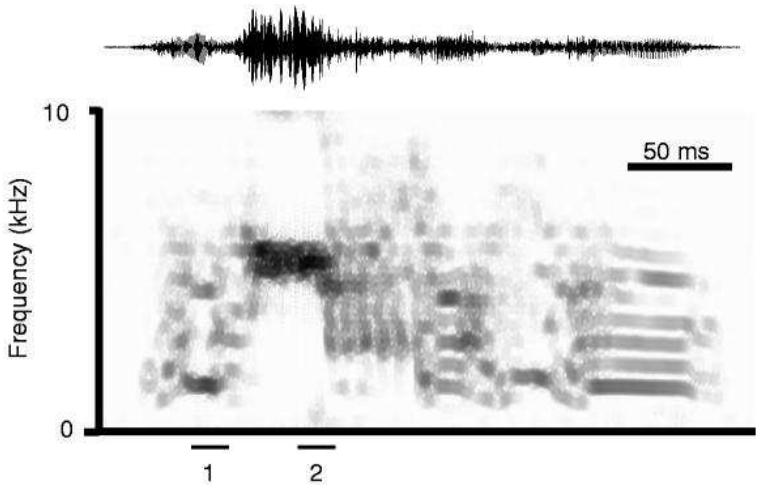
Order: Columbiformes. Vocal repertoire tonal (Gaunt et al., 1982).



sources (picture and spectrogram): `http://de.wikipedia.org/wiki/Bild:Lachduif.jpg` (photo by Coen Elemans, Wageningen Universiteit) and Elemans (2004)

Zebra finch (*Taeniopygia guttata*)

Order: Passeriformes (passerine or songbird). Vocal repertoire tonal, non-tonal periodic, non-tonal aperiodic (chaotic) (Fee et al., 1998; Fee, 2002, and references therein).



1,2 : transition from a nearly periodic signal to a noisy signal

sources (picture and spectrogram): `http://de.wikipedia.org/wiki/Bild:Zebrafinken.jpg` and Fee et al. (1998)

Table 2.3: Summary table of three different taxa of birds.

species	$\mathscr{F}_0$	spectral content	Refs.
parrots	no shift	slight effect	Brittan-Powell et al. (1997b)
pigeons	no shift	slight effect	Ballintijn and ten Cate (1998)
songbirds	no shift	appearance of harmonic overtones	Nowicki (1987)

Table 2.4: Songs recorded in heliox atmosphere of three different species of birds: qualitative changes with respect to normal vocalisations.

that a possible exhaustive explanation could be found in a source-filter model, i.e. a multi-frequency harmonic sound is filtered to a pure tone sound by vocal tract resonances (Fant, 1960). Summarizing, three hypothesis could explain pure tone birdsong production (Beckers et al., 2003a):

- a vibrating valve produces a multifrequency harmonic sound, which is filtered to a pure tone by a vocal tract resonance filter.
- vocal tract resonances are coupled to the syrinx, causing a valve structure to vibrate sinusoidally and produce a pure-tone sound.
- a yet-unknown source mechanism produces pure tone sound.

Sound signals recorded close to the syringeal sound source showed that in two closely related non-oscine species of turtledove (*Streptopelia risoria* and *Streptopelia decaocto*) pure-tone vocalizations originate through filtering of a multifrequency harmonic sound source (Beckers et al., 2003a). In ring doves (*Streptopelia risoria*), the upper esophagus, oral and nasal cavities collect expired air during vocalization (Riede et al., 2004). The level of the second and third harmonics, $2\mathscr{F}_0, 3\mathscr{F}_0$ is reduced in the inflated esophagus and emitted vocalization compared with the trachea, suggesting that the trachea and inflated esophagus act in series as acoustically separate compartments attenuating harmonics by different mechanisms (Riede et al., 2004). However, as pointed out in Ballintijn and ten Cate (1998), because of the large diversity in the morphology of the avian vocal organ, as well as the acoustic signal in birds, it can not be ruled out entirely that other hypothesis (such as the hole-tone whistle hypothesis) may play a role in some avian species.

In general, as clearly recognized by Greenewalt (1968) and re-emphasized by other authors (Gaunt and Gaunt, 1985; Fletcher, 1988), there is no simple mechanism used by all birds to produce sounds for communication. In order to explain sound generation in birds, as pointed out by Elemans et al. (2003), several qualitative models of syringeal function have been debated (Greenewalt, 1968; Gaunt and Gaunt, 1985; Goller and Larsen, 2002, and references therein). Until about 1990, only a few attempts have

been undertaken to describe the physics quantitatively (Brackenbury, 1979; Casey and Gaunt, 1985; Fletcher, 1988). The issue has been given a new impulse due to the direct endoscopic observations in birds by Goller and Larsen (1997a,b), leading to quantitative approaches to the mechanisms underlying songbird sound production based on human vocal fold models (Fee et al., 1998; Fee, 2002; Gardner et al., 2001; Laje et al., 2002; Laje and Mindlin, 2002), as discussed in the next section.

2.3.3 Syrinx models

Several models of the avian syrinx have been developed in the twentieth century (Casey and Gaunt, 1985, and references therein). However, we will be concerned here only with more recent models starting from Greenewalt (1968). We will denote as *Membrane models* the classical models, where the MTM are considered as the vibrating structure (cf. *Aero-acoustic models* in Elemans et al., 2003), whereas *Labia models* (*Modified oscillator models* in Elemans et al., 2003) will refer to more recent models, where labial masses are represented by one or two coupled nonlinear oscillators (as in many vocal fold models).

Membrane models

Greenewalt (1968) proposed a model in which both voiced and whistled songs are produced by vibration of the membranes of the syrinx. The frequency of membrane's vibration is controlled by its thickness and tension. Bernoulli forces control the motion of the membranes. The trachea acts as acoustic filter to attenuate or emphasize certain frequencies depending on the relation between the acoustic impedance of the trachea and that of the syrinx. In whistled songs, membranes vibration amplitude is assumed to be small compared to the membranes distance from the opposite wall of the trachea. Small amplitudes cause nearly sinusoidal membranes vibrations, resulting in an acoustic output with only a single frequency component (see Section 2.1.4). On the other hand, voiced sound is assumed to appear when membranes vibrational amplitude approaches closely, or even contacts, the opposing tracheal wall, resulting in a more rich-harmonic spectral content.

Brackenbury (1979) developed a model based upon theoretical discussion on a vibrating surface panel exposed to a semi-infinite uniform tangential air flow by Ffowcs Williams and Lovely (1975). This discussion was an attempt to add some quantitative detail to the general model of Greenewalt (1968). The equation of motion of the membranes displacement x reads:

$$\ddot{x} + r\dot{x} + \left(k(x) - \rho a A U^2\right) = 0 \tag{2.35}$$

where ρ is the air density, a the membranes radius, A a constant governed by local boundary conditions, U the flow in the syrinx, m, r are the membrane mass and damping, and $k(x)$ is a nonlinear stiffness which rises from a minimum resting value to a maximum when the membranes are fully stretched, by for instance, the suction forces of the airstream (Brackenbury, 1979). Nonlinear self-sustained oscillation is achieved by counteracting suction forces (due to the acceleration of air through the syrinx) and nonlinear elastic restoring force.

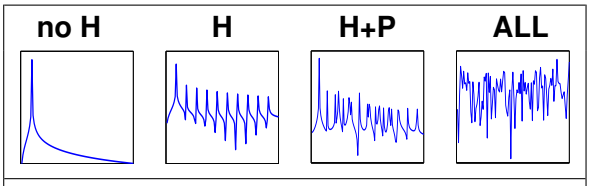
Several remarks were made by Fletcher (1988), e.g., the necessity of modifications because the membrane is acted upon by a flow constrained to lie in a tube, rather than by a semi-infinite flow field. Moreover, the membranes vibration amplitudes required to produce typical measured power levels were found to be unreasonably large (for further discussion, see Fletcher, 1988; Elemans et al., 2003).

Casey and Gaunt (1985) elaborated two different models, one for the voiced sound and one for whistled sound. The models described in Greenewalt (1968) and other precursors, (Vibrating Membrane Model - or VMM - in Casey and Gaunt, 1985) correlate with vocalizations whose spectra contain many harmonics and partials of the fundamental frequency[11], i.e., only a few of the many types of vocalizations found in birdsongs. Casey and Gaunt (1985) proposed two different models to take into account those vocalizations (widely spread among bird songs) whose spectral content is made of just the fundamental (pure tone), harmonically related frequencies or broad band noise. The first model is the so-called VSM (Vibrating String Model), which is a limiting case of the VMM: the MTMs are drawn taut into a string-like shape. The acoustic spectrum generated by the VSM contains only harmonics of the fundamental frequency and no partials.

Whistled sound, on the other hand, is thought to be produced by the so-called ADM (Aero Dynamic Model) with an entirely different mechanism, i.e., a pure whistle effect without the intervention of any vibration. Oscillations are induced by shear layer instability in air flow. With the ADM, acoustic spectra can contain the fundamental only (pure tone), harmonic overtones or broadband noise (see Table 2.5).

[11] Given a spectrum of frequencies $(\mathscr{F}_0 < \mathscr{F}_1 < \ldots < \mathscr{F}_n < \ldots)$ of an acoustic signal with fundamental frequency $\mathscr{F}_0$, we indistinctly adopt in this text the terms harmonic or overtone to indicate every frequency $\mathscr{F}_i, i > 0$. In Casey and Gaunt (1985), a more precise terminology is assumed: every frequency $\mathscr{F}_i, i > 0$ is termed overtone. Furthermore, overtones which satisfy:

- $\mathscr{F}_i = n\mathscr{F}_0, n \in \mathbb{Z}$ are called harmonic overtones (or simply harmonics)
- $\mathscr{F}_i = n\mathscr{F}_0, n \notin \mathbb{Z}$ are called partial overtones (or simply partials)



	no H	H	H+P	ALL
VMM			✓	
VSM		✓		
ADM	✓	✓		✓

Table 2.5: Harmonic content of three classical avian models (VMM, VSM and ADM): no H: fundamental and no harmonics (pure tone sound), H: fundamental and harmonics only, H+P: fundamental, harmonics and partials (these vocalizations are typically described as "noisy" in Casey and Gaunt, 1985), ALL: all frequencies present (broadband noise).

Fletcher (1988) developed the first quantitative membrane model of the avian syrinx, considering only a single bronchus and a single syringeal membrane. The model allows the calculation of syringeal membrane motion, volume flow U, tracheal pressure waveforms P and radiated acoustic power $\mathscr{P}$. The radiated power spectrum consists of harmonically related frequencies, in excellent agreement with a typical spectrum of the Australian raven (*Corvus mellori*) when anatomical parameters of this bird were used. The aerodynamics of the syringeal membranes are driven by the pressure difference $P_s - P_{\text{sup}}$[12] as a function of the acoustic impedance Z_G (representing the effective flow resistance between the air sac and the primary bronchus) and acoustic inertance of the constriction formed by the membrane.

Pressure force $F(t)$ is then calculated by means of Bernoulli equation and jet assumption. The equation of motion for the fundamental mode of vibration (other modes are treated in the same manner) reads:

$$\ddot{x} + 2r\dot{x} + \omega(x - x_0) = \frac{\varepsilon}{m} F(t) \tag{2.36}$$

where x is the maximum displacement (the central displacement for the fundamental mode), m is an effective mass associated with the mode (lower than the membrane mass), ω is the mode frequency, x_0 the equilibrium position of the membrane, and ε measures the coupling between the mode in question and $F(t)$ (however, it is normally set to unity). While excitations of more than one mode are possible, the usual result of strong nonlinearity is that one mode becomes dominant and all others are effectively driven by it (mode locking), which results in a complex but exactly repetitive waveform (Fletcher, 1978). Therefore, the spectrum of natural frequencies of the membrane is of relatively little importance for a non-linearly driven system, and there is no need for concern about this, in contrast with the VSM model of Casey and Gaunt (1985) (for

[12] We generalize here the term "-glottal" in P_s, P_{sup} to "vibrating membrane".

further discussion, see Fletcher, 1978, 1988). The collision force is modelled adding a nonlinearity to the left hand side of eq. 2.36: $r = Er$ if $x < 0$, where E is a large number of order 10 to 100. Elastic nonlinearities as in Brackenbury (1979) are omitted. A further nonlinearity is added for large vibration amplitudes by increasing the effective vibrating mass.

Raucous inharmonic screech ("screeched" song) can be easily reproduced for high air sac pressure and high membrane tension. On the other hand, an attempt to obtain whistled song is made by assuming small amplitude vibrations, accordingly to Greenewalt (1968). However, as pointed out by Fletcher (1988), the radiated sound power is quite low. Moreover, because of the weakness of the nonlinearity, the second membrane mode is only partly suppressed, and frequency components arising from it are present in the song. Therefore, the model does not fully provide a whistled sound mechanism explanation.

However, aside from the most compelling evidence that vocalizations in birds are not due to vibrations of the MTM, an alternative objection to the model presented by Fletcher comes from Laje et al. (2002): in Fletcher's model the pressure assumed to drive the MTM was composed of harmonically complex pressure waveform, in contrast with the simple pressure traces measured by Goller and Suthers in a variety of vocalizations (Laje et al., 2002, and references therein). Because the boundary conditions of the membrane are those of a vibrating drum, the difficulty of generating harmonically related overtones with a pinned membrane undermines the assumption that the source of sound is a membrane constrained like a pinned drum.

Labia models

Fee et al. (1998); Fee (2002) proposed a model which attempts to explain the role of nonlinear dynamics of the syrinx in the vocalizations of zebra finches (*Taeniopygia guttata*). The songs of this songbird reveal sudden transition from periodic to aperiodic or chaotic dynamics, period doubling, and mode-locking transitions. *In vitro* excised oscillations of zebra-finch syrinx were quite similar in oscillation frequency and amplitude to normal zebra-finch songs, suggesting that, rather than direct muscular manipulation, the origin of zebra-finch song features are due to intrinsic (nonlinear) dynamics in the syrinx. Furthermore, simple manipulations of syringeal configuration (deflection of the LL into the bronchial lumen) increased the fundamental frequency in a nonlinear fashion, exhibiting frequency jumps and oscillations which seem to be constrained to frequencies which are close to a small integer sub-harmonic of a common higher frequency (mode locking).

Stroboscopic imaging technique to visualize *in vitro* oscillations revealed that the ML, together with the attached MTM, oscillates during sound production. However,

largest motion amplitude was observed in the ML, rather than the much lighter MTM, suggesting that the ML plays a central role in syrinx function, in agreement with Goller and Larsen (1997a).

As a check of consistency, the authors developed a biophysical model of the syrinx following previous works of Ishizaka and Flanagan (1972) and Fletcher (1988). The model exhibits qualitatively similar behavior to *in vivo* zebra finch songs as well as the *in vitro* syringeal oscillations: period doubling and transitions from periodic to chaotic dynamic, as well as mode locking transitions. Possible mechanisms that might underlie mode-locking are discussed, e.g., coupling of the oscillatory mechanism to tracheal resonances.

In Fee (2002), a two-mass system is proposed as a consistency check of experimental measurements of the vibrational modes of the MVM (ML + MTM) system. Results are consistent with theoretical observations, whereby the oscillation frequency of the syrinx is determined by the dominant resonant mode of the vibrating membrane (Fletcher, 1978, 1988). The frequency of the lowest vibrational mode is determined largely by the mass of the heavier ML, rather than the thinner MTM, suggesting once more that the ML is critical in determining the oscillatory frequency of the syrinx. The MTM could play a role in the observed wave-like motion of the MVM, which has been shown to be critical for the efficient generation of sound in human larynx (Titze, 1988). Finally, possible mechanisms of frequency tuning other than changing MTM tension as in Greenewalt (1968) are discussed.

The flapping model is a model derived from Titze (1988) presented in Gardner et al. (2001), Laje et al. (2002), Laje and Mindlin (2002) and Laje and Mindlin (2005). As pointed out by Elemans et al. (2003), these models were not designed to include an accurate morphology of the syrinx, air-sacs and muscles, but were developed to understand how simple neural activation patterns (achieved by smooth variation of few parameters) can lead to complex sound signals in songbird. All models are governed by similar principles: oscillations of labia are described by relaxation oscillators, i.e., periodic systems which have fast and slow varying "configurational" states in their period (Jackson, 1989). Relaxation oscillators are typically characterized by a relatively slow storage of energy, followed by a rapid release of this energy. An example is the well-known *van der Pol* oscillator:

$$\ddot{x} - \varepsilon(1 - x^2)\dot{x} + x = 0 \tag{2.37}$$

for large ε (Berge et al., 1986, p. 28).

In Laje et al. (2002), the equation of motion describing the departure of the midpoint of the labia, x, from the pre-phonatory position reads:

$$\ddot{x} = kx - r\dot{x} - c\dot{x}x^2 - F_0 \tag{2.38}$$

where F_0 is a constant force term and c is the nonlinear dissipation term which takes place when either the labia meet each other or the containing walls.

The flapping model has been used with a slightly modified form than eq. 2.38, to show how vocalizations of the canary (*Serinus canaria*), e.g., starts, stops and paused between syllables (as well as pitch and timbre) are inherent in the mechanics and can often be expressed by simply varying smoothly two driving parameters, respectively P_s and k (Gardner et al., 2001).

In Laje et al. (2002), the flapping model described by eq. 2.38 has been tested successfully to reproduce experimental recordings of the chingolo sparrow (*Zonotrichia capensis*). Coupling with the trachea is ignored, and simulations suggest that significative coupling is not an issue. However, because a high degree of coupling is given, among the other things, by a very small $\frac{A}{L}$ ratio (vocal tract cross section to vocal tract length), and the syrinx sits deep in the thoracic chamber of songbirds, some of the complex elements of birds vocalizations could be the result of these feedback interactions.

Laje and Mindlin (2005) presented a simple model for birdsong production in Oscine songbirds, taking into account both source-source and source-tract acoustic interactions. In the former (source-source), the vocal sources are not independent from each other, in the latter (source-tract), acoustic coupling between sources and vocal tract are considered (i.e., the traditionally source-filter separation hypothesis does not hold). With this study, the flapping model shows that there might be mechanism, other than intrinsic nonlinearity of the labia, underlying complex features in spectrograms.

Finally, Mindlin et al. (2003) presented an experimental validation of the flapping model for the production of birdsongs by comparing recorded and synthetic songs, obtained by driving the model with functions whose time dependence came from recordings of muscle activities and air sac pressures. The model produced recognizable songs and suggests mechanisms by which the activities of different syringeal muscles control the avian vocal organ for song production, helping to build a bridge between the behavioral output (song) and the complex representation and execution of vocal communication at the level of the brain (for details, see Laje et al., 2002; Laje and Mindlin, 2002; Mindlin et al., 2003).

3

Modelling Bird Songs: Voice Onset, Overtones and Registers

This chapter is an extended version of Zaccarelli, R., Elemans, C. P. H., Fitch, W. T. and Herzel, H. (2005), "Two-Mass Models of the Bird Syrinx", Proceed. 4th MAVEBA Workshop (Zaccarelli et al., 2005) and Zaccarelli, R., Elemans, C. P. H., Fitch, W. T. and Herzel, H. (2006), "Modelling Bird Songs: Voice Onset, Overtones and Registers", Acta Acustica united with Acustica (Zaccarelli et al., 2006).

We analyze two symmetric two-mass models of the avian syrinx. Our first model applies to songbirds and is a rescaled version of the well-known human two-mass model. Our second model (trapezoidal model) introduces a smoother geometry and is used to simulate the ring dove (*Streptopelia risoria*) syrinx. Simulations show that both models exhibit self-sustained vibrations. We show that the occurrence of collisions and the intensity of harmonics depend strongly on the configuration of the syrinx. The songbird model does not present instabilities. The trapezoidal model, however, displays coexisting limit-cycles that represent vibrations with, and without collisions at the same pressure. Register-like transitions are accompanied by subharmonics and deterministic chaos.

3.1 Introduction

Two-mass models of mammalian vocal fold vibration have been used successfully to describe the normal voice (Ishizaka and Flanagan, 1972; Pelorson et al., 1994), vocal fold paralysis (Isshiki et al., 1978; Smith et al., 1992; Steinecke and Herzel, 1995; Mergell et al., 2000), prostheses design (Lous et al., 1998), phonation onset (Mergell et al., 1998), voice instabilities at high pressures (Jiang et al., 2001) and source-tract coupling (Mergell and Herzel, 1997; Hatzikirou et al., 2006).

As introduced in Section 2.3, in contrast to most mammals, birds do not generate vocalizations with their larynx but with their unique vocal organ, the syrinx (Greenewalt, 1968). Instead of vocal folds, thickened membranes called "labia" serve as vibrating tissue (Goller and Larsen, 1997a; Fee et al., 1998; Larsen and Goller, 1999, 2002). However, sound production in birds is thought to be based on aerodynamical principles similar to that of human phonation (Goller and Larsen, 1997a; Fee et al., 1998; Goller and Larsen, 2002; Mindlin and Laje, 2005). Consequently, similar modelling approaches might be applicable and several types of biomechanical models have been developed (Fee et al., 1998; Fletcher and Tarnopolsky, 1998; Gardner et al., 2001; Fee, 2002; Laje et al., 2002; Elemans et al., 2003; Mindlin and Laje, 2005). It is not obvious that rescaling of the original two-mass model will lead to appropriate oscillations at realistic driving pressures and damping ratios. The syrinx is generally much smaller than the human larynx, which leads to smaller areas for the interactions between airflow and vibrating structures and to higher fundamental frequencies of the produced sound.

In this section, we develop two biomechanical models of the syrinx to study the onset of sound generation and control of higher harmonics (overtones) in the absence of source-tract coupling.

Low-order models, as discussed in this paper, are over-simplifications of the physiology. More realistic geometries (Pelorson et al., 1994), additional degrees of freedom (Story and Titze, 1995), coupling to the resonators (Lous et al., 1998) or a more detailed description of the jet separation (Pelorson et al., 1994) can improve the simulated sound signal significantly. In this paper we consider rather simple models with somewhat less realistic output. These core models allow comprehensive bifurcation analyses and we can address basic questions: Do we find reasonable vibrations of the masses even for much smaller geometries? How do symmetric upper and lower masses and a trapezoidal shape influence voice onset, spectral slope and voice instabilities?

In a first model version (rescaled two-mass model), we adapt the well-known simplified two-mass model to the dimensions of a songbird syrinx. Songbirds have two

pairs of bilateral labia, which can operate as two independent sound sources (Suthers, 1990; Suthers et al., 1999). Our model considers one unilateral pair of labia.

In the second model (trapezoidal model), a more realistic geometry of the so-called LTM (Lateral Tympaniform Membrane) is adopted to describe the syrinx of a non-songbird: the ring dove (*Streptopelia risoria*). The upper and lower masses are connected with mass-less plates on which pressure can act. Such a configuration with massless plates has been applied to reproduce experimental data from the human voice and to design laryngeal prostheses (Lous et al., 1998).

Both models exhibit self-sustained oscillations at physiologically realistic parameter values. In the classical two-mass model, as well as in the model of the songbird syrinx, collisions occur at medium pressures leading to strong harmonics. In the model of the dove syrinx, however, collision is partially avoided, leading to more pure tones. We relate these observations to the widely discussed topic of how birds control the intensity of their harmonics (Klatt and Stefanski, 1974; Nowicki, 1987; Williams et al., 1989; Nowicki et al., 1992; Beckers et al., 2003a; Riede et al., 2004; Beckers et al., 2004; Fletcher et al., 2004). The trapezoidal model exhibits coexisting vibratory regimes that resemble vocal registers. At the same subsyringeal pressure, vibrations with and without collisions are possible. Slow variation of subsyringeal pressure can induce subharmonics, deterministic chaos and a sudden jump to the other "registers".

3.1.1 The Models: Overview

Both models are governed by the same equations of motion:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = v_1 \tag{3.1}$$

$$\frac{\mathrm{d}v_1}{\mathrm{d}t} = \frac{1}{m_1}\Big(F_1 - r_1 v_1 - k_1 x_1 + C_1 - k_c(x_1 - x_2)\Big) \tag{3.2}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = v_2 \tag{3.3}$$

$$\frac{\mathrm{d}v_2}{\mathrm{d}t} = \frac{1}{m_2}\Big(F_2 - r_2 v_2 - k_2 x_2 + C_2 - k_c(x_2 - x_1)\Big) \tag{3.4}$$

where F_i are the pressure forces derived from the Bernoulli equation and the jet assumption and C_i represent the collision forces. We will give a detailed description of both models in the next sections (for technical details of the trapezoidal model, see Appendix A).

3.2 Rescaled two-mass model

3.2.1 Derivation and voice onset

The classical two-mass model directly describes the well-known phase shift between upper and lower edge of aerodynamically driven vibrating tissue (Titze, 1988, 1994). Originally, this model was derived in order to reproduce human vocal fold vibrations (Ishizaka and Flanagan, 1972; Pelorson et al., 1994; Lous et al., 1998; Sciamarella and D'Alessandro, 2004; Hatzikirou et al., 2006). Therefore, it is not obvious that a rescaling of the original two-mass model will lead to appropriate oscillations at realistic driving pressures and damping ratios.

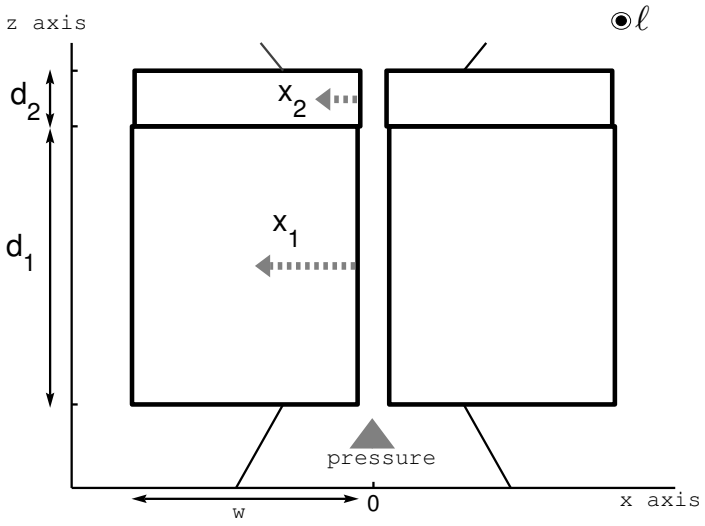


Figure 3.1: The rescaled two-mass model of the songbird syrinx.

3.2.2 Parameters setup

The fundamental frequency $\mathscr{F}_0$ of many bird songs is on the order of 1 kHz. If we assume similar tissue elasticities (k_1, k_2) and density ρ as in human vocal fold modelling (Ishizaka and Flanagan, 1972), we can derive the appropriate parameters of the rescaled model shown in Table 3.1. We maintain the 5:1 ratio of lower to upper mass thickness (compare Fig. 3.1).

The pressures are derived in earlier studies (Steinecke and Herzel, 1995) from the

symbol	description	value
ℓ	length of the syringeal lumen	0.3 cm
a_{01}	lower rest area	0.0021 cm^2
a_{02}	upper rest area	0.00175 cm^2
d_1	1st mass thickness	0.1 cm
d_2	2nd mass thickness	0.02 cm
m_1	1st mass	0.0015 g
m_2	2nd mass	0.0003 g
k_1	1st mass stiffness	0.08 g/ms^2
k_2	2nd mass stiffness	0.008 g/ms^2
r	damping constant ($r_1 = r_2$)	0.002 g/ms
k_c	coupling constant	0.025 g/ms^2

Table 3.1: Parameters of the rescaled two-mass model shown in Fig. 3.1.

Bernoulli equation using a jet assumption (see Section 2.1.2):

$$P_1 = P_s \left(1 - \Theta(a_{min}) \left(\frac{a_{min}}{a_1}\right)^2\right) \Theta(a_1) \tag{3.5}$$

$$P_2 = 0 \tag{3.6}$$

where $\Theta(x)$ is the *Heaviside* function:

$$\Theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ \\ 0 & \text{if } x \leq 0 \end{cases}$$

The forces F_i, C_i read as in previous studies (Steinecke and Herzel, 1995):

$$F_1 = \ell d_1 P_1 \tag{3.7}$$

$$F_2 = 0 \tag{3.8}$$

$$C_i = -\Theta(-a_i) c_i \frac{a_i}{2\ell}, \quad i \in \{1,2\} \tag{3.9}$$

Viscous resistance of the vibrating tissues can be expressed in terms of damping ratio $\zeta = \frac{r}{2\sqrt{km}}$ (Ishizaka and Flanagan, 1972). Because of the decreasing mass m, r was rescaled to keep approximately the same ζ values as in (Ishizaka and Flanagan, 1972). There is almost no information available on the prephonatory syringeal width (rest areas values) and on the prephonatory shape (expressed by the ratio $\frac{a_{02}}{a_{01}}$) of vibrating tissues in avian phonation studies. Therefore, we have chosen rest areas a_{01} and a_{02} that allow easy vibrations, i.e., a low onset pressure (according to the values reported, e.g., in Ishizaka and Flanagan, 1972). The importance of the prephonatory glottal width for the onset of oscillations has been investigated by Titze (1988) for a rectangular glottis (i.e.,

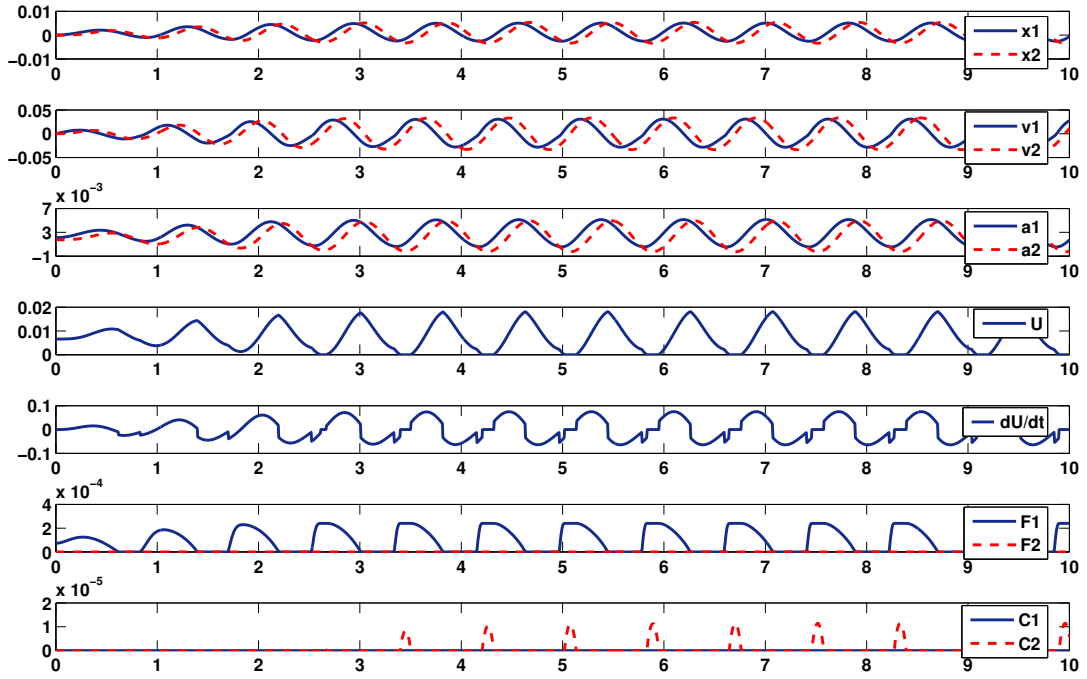


Figure 3.2: Simulation results of the rescaled two-mass model with subglottal pressure $P_s = 8$ cm H_2O. x axis scale: time (msec), y axis scale according to the units given in Table 3.1 (and their relative combinations).

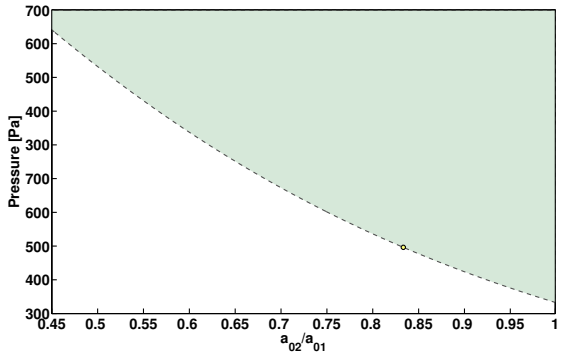


Figure 3.3: Onset pressure versus the ratio $\frac{a_{02}}{a_{01}}$. The shaded area above the bifurcation line denotes the region where self-sustained oscillations occur. The dot represents the default value of the rescaled two-mass model.

assuming $a_{01} = a_{02} = a_0$): the closer the vocal folds are brought together, the easier is to initiate small-amplitude oscillation. Consistent with this assumption, in the rescaled two-mass model the driving pressure $P_1 \simeq 0$ for too large values of a_0 (compare eq. 3.5. We remark that $a_i = a_0 + 2\ell x_i$ by definition). Furthermore, systematic variations of the configuration revealed that a rectangular or a slightly convergent shape allow realistic onset pressures (Fig. 3.2 and 3.3).

Fig. 3.4 shows the onset of self-sustained oscillations (solid Hopf bifurcation line)

for increasing subsyringeal pressure P_s and varying stiffness k_1. The onset pressure of the rescaled model can be below 0.004 $\frac{\text{g}}{\text{cm ms}^2}$ (400 Pascal $\approx$ 4 $\text{cm}\,\text{H}_2\text{O}$[1]) around our default parameters. Thus, rescaled dimensions with proper damping ratios and rest areas lead to an onset of oscillations at realistic pressure values. Therefore our rescaled two-mass model can serve as a first step to model vibrating tissues in the syrinx.

3.2.3 Intensity of overtones

The intensity of higher harmonics (overtones) is a widely discussed topic in bird song studies (Klatt and Stefanski, 1974; Nowicki, 1987; Williams et al., 1989; Nowicki et al., 1992; Beckers et al., 2003a, 2004; Riede et al., 2004; Fletcher et al., 2004). A pure tone (e.g., a sine wave) has no overtones (see Section 2.1.4). Collisions of the vibrating tissues, however, lead to pronounced harmonics (Titze, 1994). To study the intensity of overtones in the rescaled two-mass model, we calculate the power spectrum of the flow derivative $\frac{\text{d}U}{\text{d}t} = \dot{U}$ (we recall from Section 2.1.4 that $\dot{U}$ is a reasonable approximation of the radiated sound pressure, see Ishizaka and Flanagan, 1972; Rothenberg, 1981; Titze, 1994)[2]. We introduce a simple measure of the intensity of overtones - the *Harmonics Ratio* (HR):

$$HR = 10\log(\frac{H_1}{H_0})$$

which is calculated from the spectral power H_0 at the fundamental frequency and the power at the first harmonic, H_1 (see Section 2.1.4). This quantity is closely related to the widely used spectral slope (Titze, 1994; Sciamarella and D'Alessandro, 2005). Values below -20 dB indicate that the signal has weak overtones (Goller and Larsen,

[1] All units are given in centimeters, grams and milliseconds and their corresponding combinations: hence pressure is measured in $\frac{\text{g}}{\text{cm ms}^2} = 10^5$ Pa. Consequently, with

$$1\ \text{cm}\,\text{H}_2\text{O} \approx 10^2 Pa$$

we get

$$1\ \text{cm}\,\text{H}_2\text{O} \approx 0.001 \frac{\text{g}}{\text{cm ms}^2}.$$

[2] The equation of the flow in the rescaled two-mass model reads (Steinecke and Herzel, 1995):

$$U = \left(\sqrt{\frac{2P_s}{\rho}}\right) a_{min}\Theta(\max(0, a_{min})).$$

Consequently, $\dot{U} = 0$ if $a_{min} \leq 0$. If $a_{min} > 0$:

$$\dot{U} = \frac{\text{d}U}{\text{d}t} = 2\ell\sqrt{\frac{2P_s}{\rho}}\left(\Theta(a_2 - a_1)\dot{x}_1 + \Theta(a_1 - a_2)\dot{x}_2\right).$$

2002).

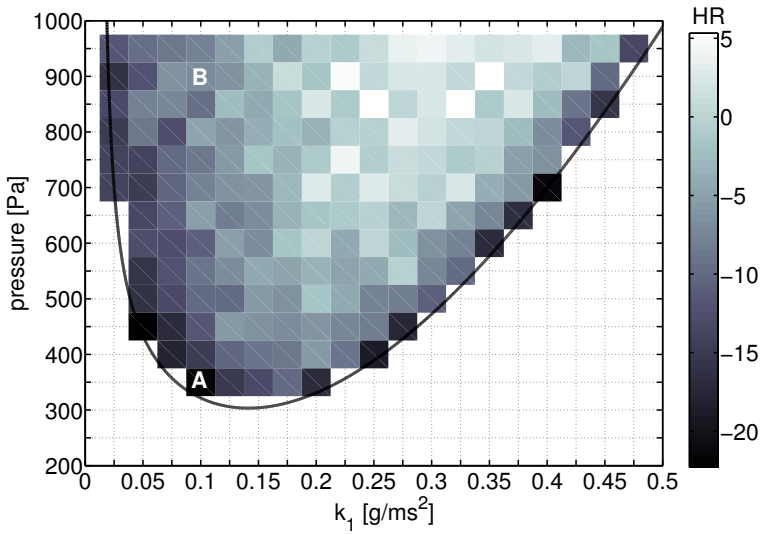


Figure 3.4: Variation of the onset pressure depending on the stiffness (solid Hopf bifurcation line) and Harmonic Ratio (HR) values color map above the onset of the pressure. At points A and B, i.e. close and relatively far away from the Hopf bifurcation, we evaluated the power spectrum of the flow derivative (see Fig. 3.5). The fundamental frequencies at Ps=900 Pa range from approximately 800 Hz (k1 = 0.0025) to 2,400 Hz (k1=0.45).

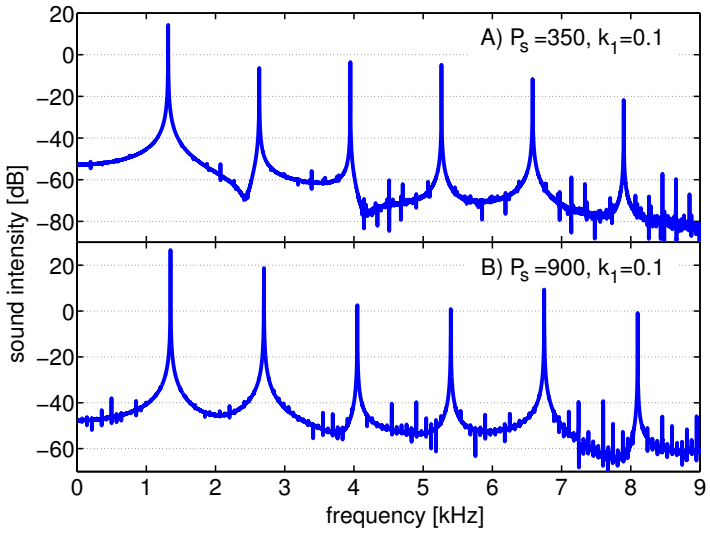


Figure 3.5: Power spectra at two different regimes corresponding to the letters A (upper panel) and B (lower panel, HR $\simeq$ -8) in Fig. 3.4. Close to the Hopf bifurcation point (A) (Ps = 0.0035) we observe less intense harmonics (HR $\simeq$ -21).

Fig. 3.4 shows the values of HR (in grey scale) for varying subsyringeal pressure P_s and stiffness k_1. Only in the immediate neighborhood of phonation onset nearly sinusoidal oscillations occur (e.g. at point A in Fig. 3.4). Point B in Fig. 3.4 represents a more typical situation: collisions of the masses lead to rather strong harmonics (Fig.

3.5). In summary, the rescaled two-mass model can be used to model sound production with pronounced overtones. Extensive exploration of the effects of parameter variations revealed that the observed periodic vibrations with collision are quite robust. No register transitions or nonlinear phenomena such as subharmonics were found (see Fig. 3.6).

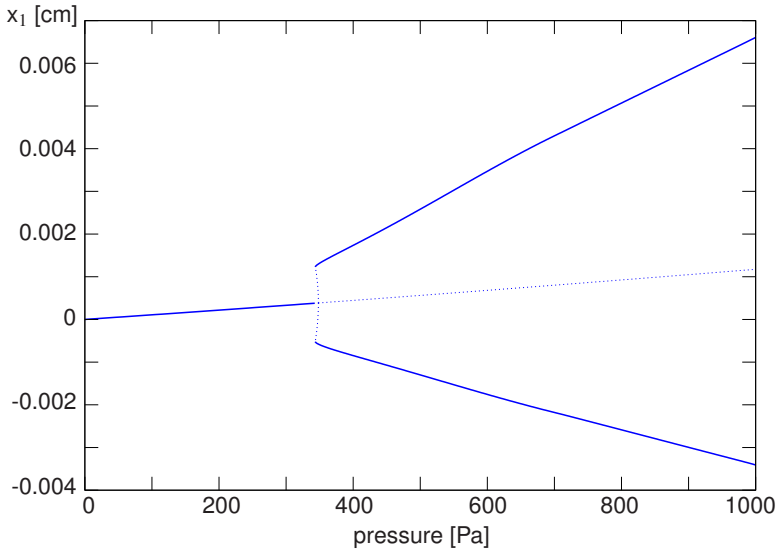


Figure 3.6: Bifurcation diagram of the displacement of the first mass, x_1, for increasing subglottal pressure P_s. We observe a subcritical Hopf bifurcation (dotted lines: unstable fixed point or limit cycles, solid lines: stable fixed points or limit cycles). No register transitions or nonlinear phenomena such as subharmonics were found (compare Fig. 3.10).

3.3 Trapezoidal model

3.3.1 Geometrical aspects

In this section we will relax some over-simplifications of the original two-mass model along the lines of Pelorson et al. (1994) and Lous et al. (1998). The trapezoidal model aims to describe the syrinx of the ring dove (*Streptopelia risoria*). In contrast with songbirds, the vocal organ of ring doves is located at the bronchotracheal junction (Gaunt et al., 1982; Goller and Larsen, 1997b), i.e. above the bifurcation of the trachea into the bronchi (see Section 2.3.1). The anatomy of the ring dove syrinx (see Fig. 3.7) suggests that a smoother model is more appropriate (*Streptopelia decaocto*, Ballintijn et al., 1995; *Streptopelia risoria*, C.P.H. Elemans unpublished results). Therefore, each side of the LTM is modelled as a system of two masses linked together by three massless

plates (see Fig. 3.8), as presented in Lous et al. (1998). The parameters listed in Table 3.2 were obtained from anatomical studies (published in Elemans et al., 2007. See Table 4.1 on p. 76). Instead of the estimated total mass of approximately 9 mg, we assume a vibrating mass of 2 mg in order to achieve a reasonable fundamental frequency. Two important modifications of the two-mass model are introduced:

- symmetry between upper and lower masses (i.e. $m_1 = m_2, k_1 = k_2, r_1 = r_2$) as in Lous et al. (1998), because there is no anatomical reason to introduce a small upper mass as in the classical two-mass model
- smoothed geometry via lower and upper plates which are characterized by the height parameters d_1 and d_3, respectively.

These modifications affect the calculation of pressure and collision forces significantly. For example, even for a closed syrinx there is a pressure force acting on the lower mass via the lower plate.

symbol	description	value
$2w$	width of the trachea	0.3 cm
ℓ	length of the trachea	0.3 cm
a_{01}	lower rest area	0.003 cm^2
a_{02}	upper rest area	0.003 cm^2
d_1	1$^{\text{st}}$ mass height	0.04 cm
$d_1 + d_2$	2$^{\text{nd}}$ mass height	0.24 cm
$d_1 + d_2 + d_3$	LTM height	0.28 cm
m	masses ($m_1 = m_2$)	0.001 g
k	stiffness	0.02 g/ms^2
r	damping constant	0.001 g/ms
k_c	coupling constant	0.005 g/ms^2

Table 3.2: Parameters of the model of the ring dove syrinx (see Fig. 3.8).

3.3.2 Calculation of forces

In the traditional two-mass model, the area exposed to pressure (or collision) is always rectangular and normal to the motion of the masses (Sciamarella and D'Alessandro, 2004). In order to derive the pressure forces F_1, F_2, we need to multiply the pressure with the corresponding area. The collision forces C_1, C_2 can be written as a linear function of the areas a_1, a_2, as in Steinecke and Herzel (1995).

In our trapezoidal model, most of the quantities have to be generalized to adapt to the new model geometry, as described in more detail in Appendix A. First, we assume

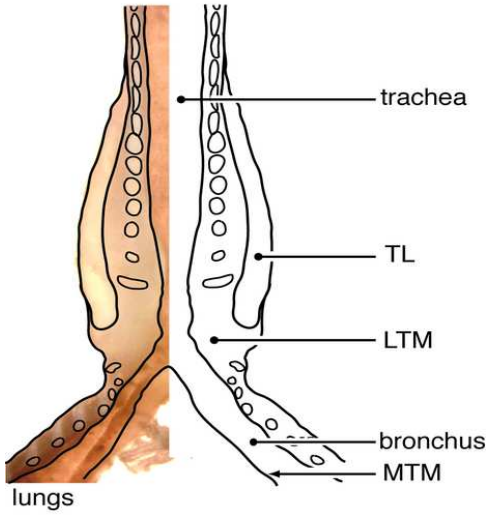


Figure 3.7: The syrinx of a ring dove (*Streptopelia risoria*). LTM: Lateral Tympaniform Membranes, MTM: Medial Tympaniform Membrane, TL: Tracheolateralis muscle.

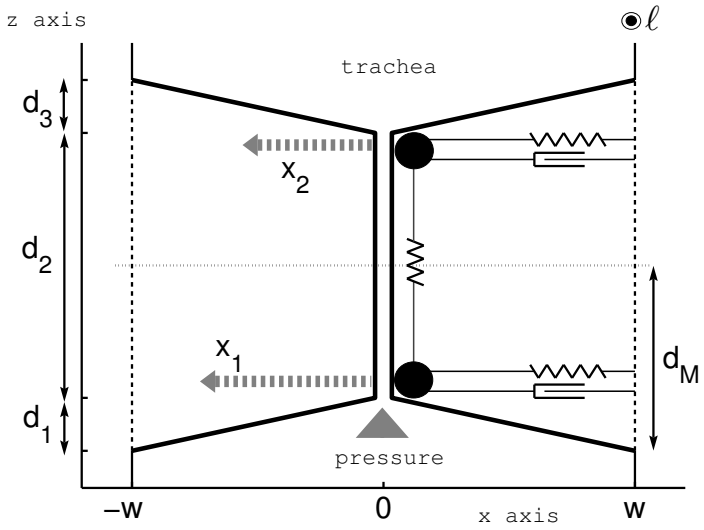


Figure 3.8: The model of the ring dove syrinx: the point masses are joined by three massless plates.

for simplicity that all forces above the imaginary horizontal midline at $d_M = d_1 + \frac{d_2}{2}$ act on the upper mass whereas all forces below that midline drive the lower mass (midline assumption). Note that in Lous et al. (1998), F_1 and F_2 are derived from a balance between forces and torques leading to a more detailed derivation of the force terms.

Second, we define the syringeal area at height z as $a(z) = 2\ell x(z)$, where $x(z)$ is obtained by means of plate equations, i.e. the linear equations which identify each plate on the plane xz. Particular values of $a(z)$ are a_1, a_2 (syringeal areas at 1^{st} and 2^{nd} mass heights, d_1 and $d_1 + d_2$). The minimum syringeal area a_{min} is defined as $a_{min} = \min\{a_1, a_2\}$.

Pressure force By means of the Bernoulli equation, jet separation assumption and $a(z)$ defined above, we can calculate the pressure $P = P(z)$ in the syringeal lumen at height z (see Appendix A):

$$P(z) = \begin{cases} P_s \left[1 - \left(\dfrac{a_{min}}{a(z)} \right)^2 \right] \Theta(z_m - z), & \text{if } a_{min} > 0 \\ \\ P_s \Theta(\zeta_{min} - z), & \text{if } a_{min} \leq 0 \end{cases}$$

where z_m is the ordinate at which a_{min} is found and, in case of collision ($a_{min} \leq 0$), ζ_{min} is the minimum collision ordinate, i.e. the minimum ordinate z for which $a(z) \leq 0$.

If z_0 and $z_1 (z_0 < z_1)$ are two generic ordinates of points belonging to the same plate and α the angle formed by that plate with the z axis, the pressure force acting on the plate area between z_0 and z_1 will be by definition $F(z_0, z_1) = \int_A P(z)\,\mathrm{d}A$, where A is the plate area

$$A = \frac{z_1 - z_0}{\cos\alpha} \ell$$

Because no vertical motion is supposed, the component along the x axis of this pressure force reads:

$$\begin{aligned} F_x(z_0, z_1) &= \cos\alpha\, F(z_0, z_1) \\ &= \cos\alpha \int_0^{\frac{z_1 - z_0}{\cos\alpha}\ell} P(z)\,\mathrm{d}A \\ &= \ell \int_{z_0}^{z_1} P(z)\,\mathrm{d}z \end{aligned}$$

Consequently, and by means of the midline assumption, the forces F_1, F_2 read:

$$F_1 = F_x(0, d_1) + F_x(d_1, d_M) \tag{3.10}$$

$$F_2 = F_x(d_M, d_2) \tag{3.11}$$

Collision force Traditional two-mass models do not require the calculation of contact area, because the projected area is rectangular and there is no gradation in opening and closing (Sciamarella and D'Alessandro, 2004). We define a collision force that is consistent with eq. 3.9 and admits a gradual variation of contact area in time. First, we remark that each collision force C_i is zero if:

$$C_i = 0 \Leftrightarrow a_i \geq 0 \text{ AND } a_M \geq 0, \quad i \in \{1, 2\}$$

If the collision force C_i is nonzero, we define respectively:

$$C_1 = \frac{1}{\delta(0, d_M)} \int_0^{d_M} -\Theta(-a(z)) a(z) \frac{c_i}{2\ell}\,\mathrm{d}z \tag{3.12}$$

$$C_2 = \frac{1}{\delta(d_M, d_3)} \int_{d_M}^{d_3} -\Theta(-a(z)) a(z) \frac{c_i}{2\ell}\,\mathrm{d}z \tag{3.13}$$

where $\delta(z_1, z_2)$ is the distance (on the line $x = 0$) between z_1 and z_2 along which $a(z) < 0$.

We chose to normalize C_1 and C_2 using distance δ in order to keep c_i of the dimension of g/ms^2 and to obtain a generalization of the previous collision force. The latter is found comparing it with the degenerate case $d_1 = 0$. For the implementation of C_i, we take into account that the integrals in eq. 3.12 and 3.13 are proportional either to triangular or trapezoidal areas of colliding masses (see Appendix A for details).

3.3.3 Bifurcations in the trapezoidal model

As in Section 3.2.3, we study the onset of oscillations and the strength of harmonics in our trapezoidal model. The upper panel in Fig. 3.9 shows that oscillations can be obtained at fairly low subsyringeal pressures (> 200 Pa). Increasing stiffness leads to an almost linear increase of the onset pressure as observed earlier in other models (Steinecke and Herzel, 1995; Mergell et al., 1999). An essential modification of the standard two-mass model is the smoother geometry, i.e. a non-vanishing height d_1. It turns out that in a certain range around our default parameter ($d_1 = 0.004$ cm) oscillations are easily obtained (lower graph in Fig. 3.9).

Fig. 3.10 shows a one-dimensional bifurcation diagram for increasing subsyringeal pressure P_s. Fig. 3.11 shows a more detailed bifurcation diagram near the onset visualizing coexisting limit cycles and negative values of the areas corresponding to colliding tissues. We observe in Fig. 3.10 a sudden onset of oscillations around 300 Pa via a subcritical Hopf bifurcation. The amplitude of the resulting limit cycle increases and no collision occurs (see Fig. 3.11). At 425 Pa another limit cycle with larger amplitudes is observed exhibiting collisions of the upper part described by x_2 (corresponding to negative values of a_2 in Fig. 3.11). Furthermore, Fig. 3.10 implies that quite distinct vibration patterns coexist at the same pressure. The large limit cycle contains stronger harmonics than the smaller one (Fig. 3.12). These observations resemble observations in previous experiments of the chest to falsetto transitions in excised larynx studies (Berry et al., 1996). Several register-like transitions were also found in simulations of an extended two-mass model (Sciamarella and D'Alessandro, 2004). Because the suppression of collision and the register transitions are novel features of our trapezoidal model, we discuss these phenomena in some detail[3].

[3] A more detailed bifurcation analysis of the trapezoidal model can be found in Appendix B. Results will be employed in the derivation of muscular functions affecting the ring dove coo (see Chapter 4).

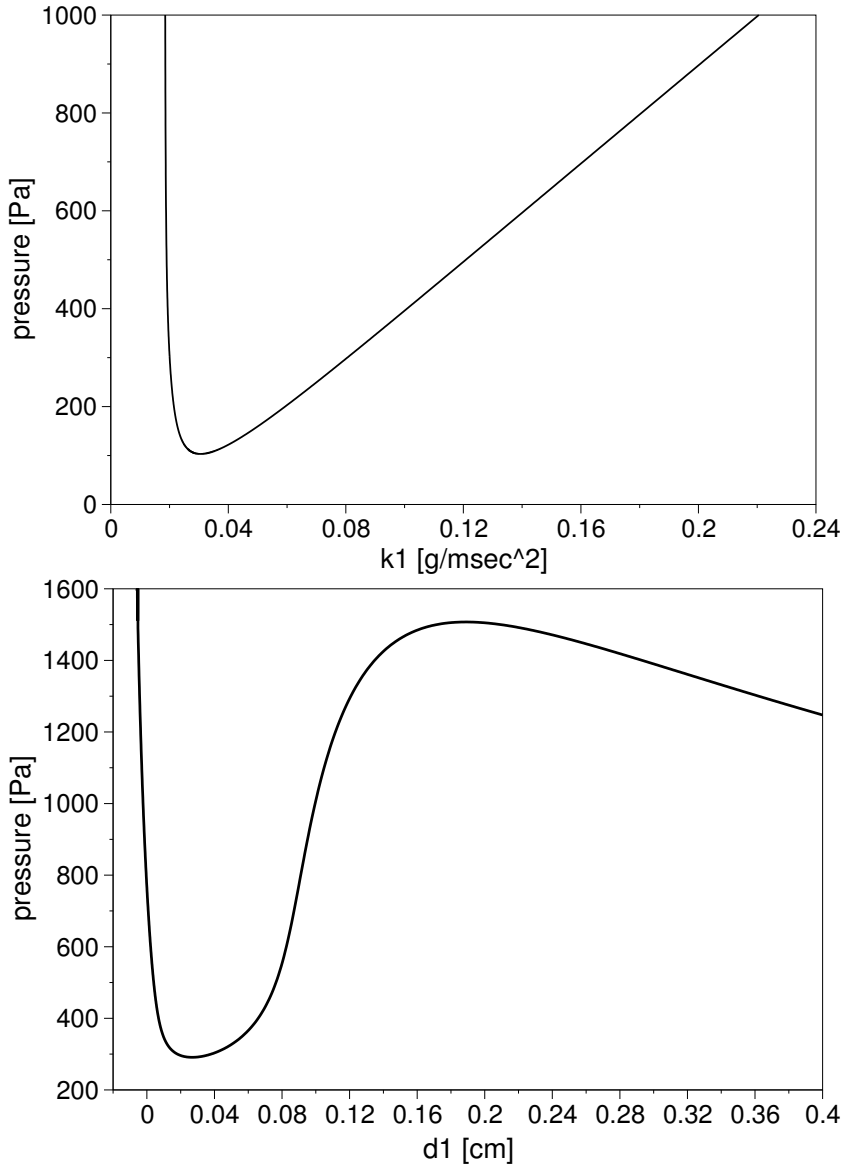


Figure 3.9: Onset pressure as a function of stiffness (upper graph) and of height d_1 (lower graph) for the trapezoidal model. Initial conditions: $[x_1, v_1, x_2, v_2] = [0, 0, 0, 0]$.

3.3.4 Avoidance of collisions

As illustrated in Fig. 3.4, the standard two-mass model is characterized by strong harmonics due to collision even at small and medium pressures. However, the lower limit cycle of our trapezoidal model is collision-free even at high pressures and consequently has only weak harmonics (lower panel in Fig. 3.12). This is due to the increasing steady state area (see Fig. 3.11) which pulls apart the masses. The equilibrium in the standard two-mass model, however, is constant, i.e. it does not depend on the subglottal pressure. If we set the derivatives with respect to time in the equations 3.1 - 3.4 to zero, we obtain in the two-mass model $x_2 = \frac{k_c}{k_2+k_c} x_1$. For a rectangular shape ($a_{01} = a_{02} > 0$) we get

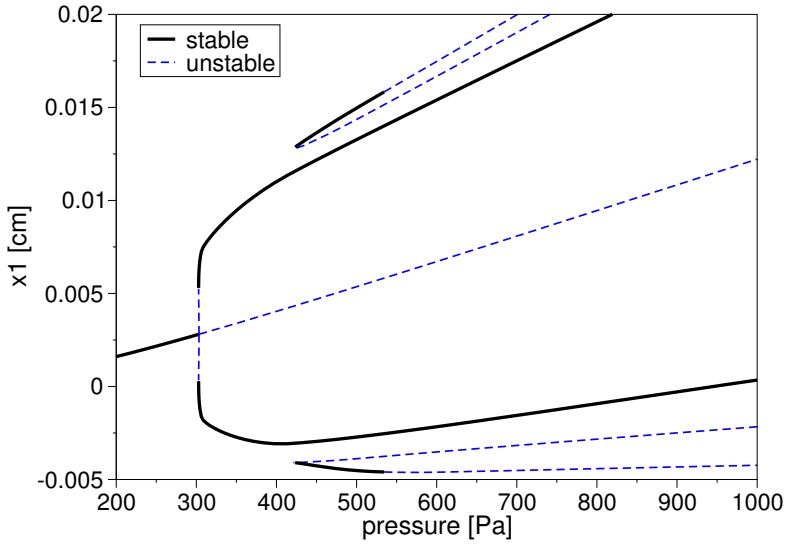


Figure 3.10: Bifurcation diagram of the variable x_1 for increasing subsyringeal pressure P_s. Note the coexistence of stable limit cycles around 450 Pa.

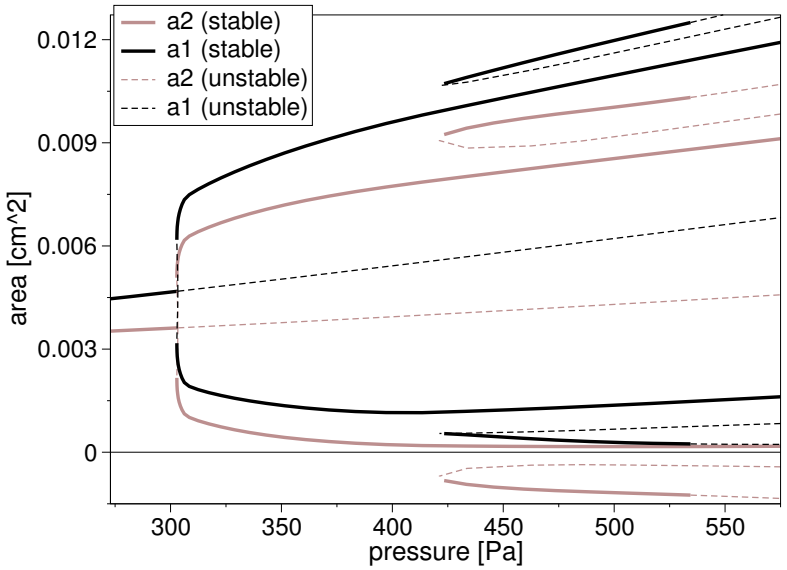


Figure 3.11: Detailed bifurcation diagram of the syringeal areas a_1, a_2 for increasing subsyringeal pressure P_s.

from eq. 3.2 the trivial solution $x_1 = x_2 = 0$. As shown in Steinecke and Herzel (1995), oscillations starting from that rest position lead to collision, even for small pressures.

If we apply the same calculations to our trapezoidal model we obtain:

$$x_2 = \frac{k_c}{k_2 + k_c} x_1 + \frac{F_2}{k_2 + k_c} \tag{3.14}$$

leading directly to

$$x_1 = \frac{(k_2 + k_c)F_1 + k_c F_2}{k_1 k_2 + k_1 k_c + k_2 k_c} > 0 \tag{3.15}$$

There is no simple analytical solution of this equation. However, because $F_1 > 0$ we get

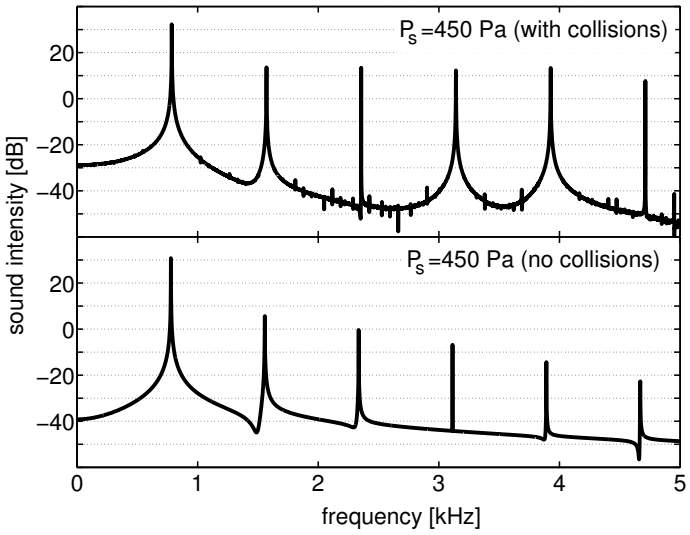


Figure 3.12: Power spectra at two different regimes: Collisions at larger vibrations lead to strong harmonics whereas harmonics decay rapidly for the small limit cycle.

no equilibrium at $x_1 = x_2 = 0$. Therefore, the smoother geometry implies that there is always a force that pulls apart the masses, and the rest position increases linearly with the subsyringeal pressure (see Fig. 3.10). Fig. 3.11 shows that for the equilibria we always have $a_1 > a_2$, i.e. a convergent configuration. Because for a convergent shape the suprasyringeal pressure has little effect on the masses, such a persistent convergent configuration might lead to a reduced source-tract interaction.

It is clearly visible that both areas remain positive for the small limit cycle, while the large limit cycle exhibits negative values of a_2 corresponding to colliding tissues.

3.3.5 Register transitions via subharmonics and chaos

As described above, we found coexistence of a small and a large limit cycle in the range of 425-535 Pa. This implies that different initial conditions lead to distinct vibration patterns.

Furthermore, small perturbations can induce sudden jumps from one attractor to another. In this section we analyze the transition from the large attractor with collisions to the small limit cycle due to a slow increase of the subsyringeal pressure. Fig. 3.13 shows the spectrogram of the sound pressure generated during the transition from the large limit cycle (with collisions) to the small limit cycle caused by a slow increase of the subsyringeal pressure. It is evident that there is a jump from harmonic rich spectra to a more sinusoidal oscillation.

Furthermore, subharmonics and noise-like components are visible. In Fig. 3.14,

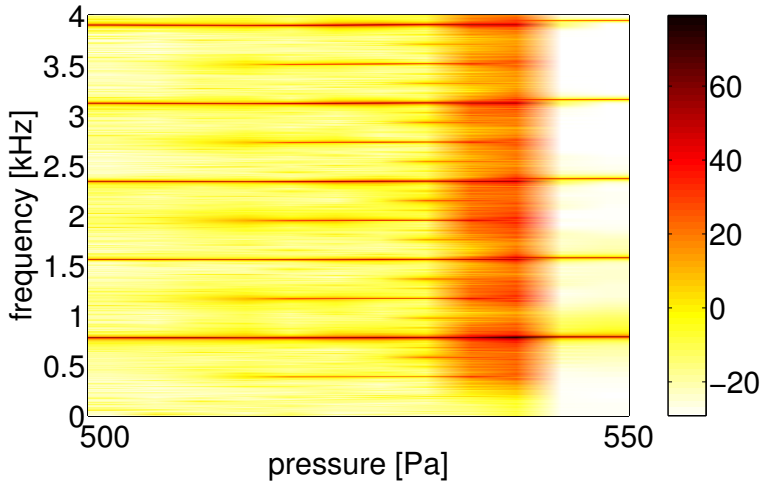


Figure 3.13: Register transitions from the large limit cycle to the small limit cycle (see Fig. 3.10). Starting from a subsyringeal pressure of 500 Pa, we observe period doubling and chaos in the vicinity of the jump to the small limit cycle with less pronounced harmonics.

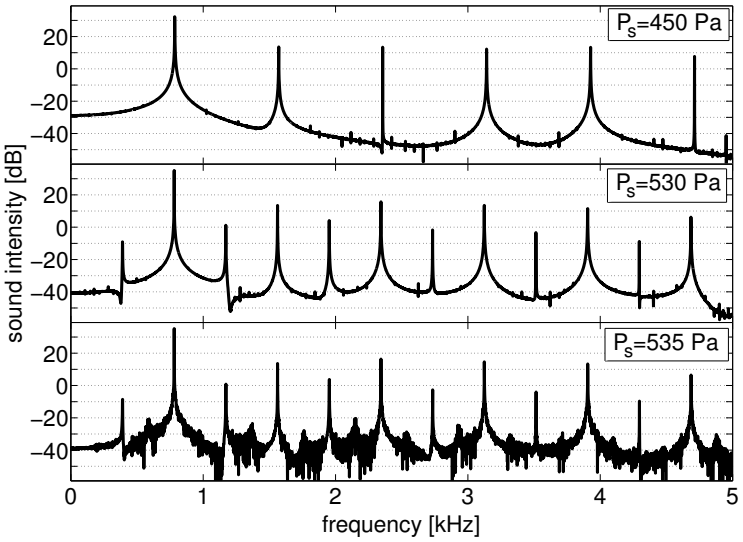


Figure 3.14: Power spectra at different pressure values for the large limit cycle depicted in Fig. 3.10: As visible from the spectrogram in Fig. 3.13, we detect a period doubling and chaos in the vicinity ($\simeq$ 540 Pascal) of the abrupt jump to the small limit cycle.

high-resolution spectra[4] confirm the appearance of subharmonics and chaos. Using phase portraits and Poincaré sections (Fig. 3.15), we have confirmed the existence of deterministic chaos in our trapezoidal model.

[4]Integration time from 1000 to 2000 ms in order to avoid transients and time step of 0.0005 msec (2000 kHz sampling rate).

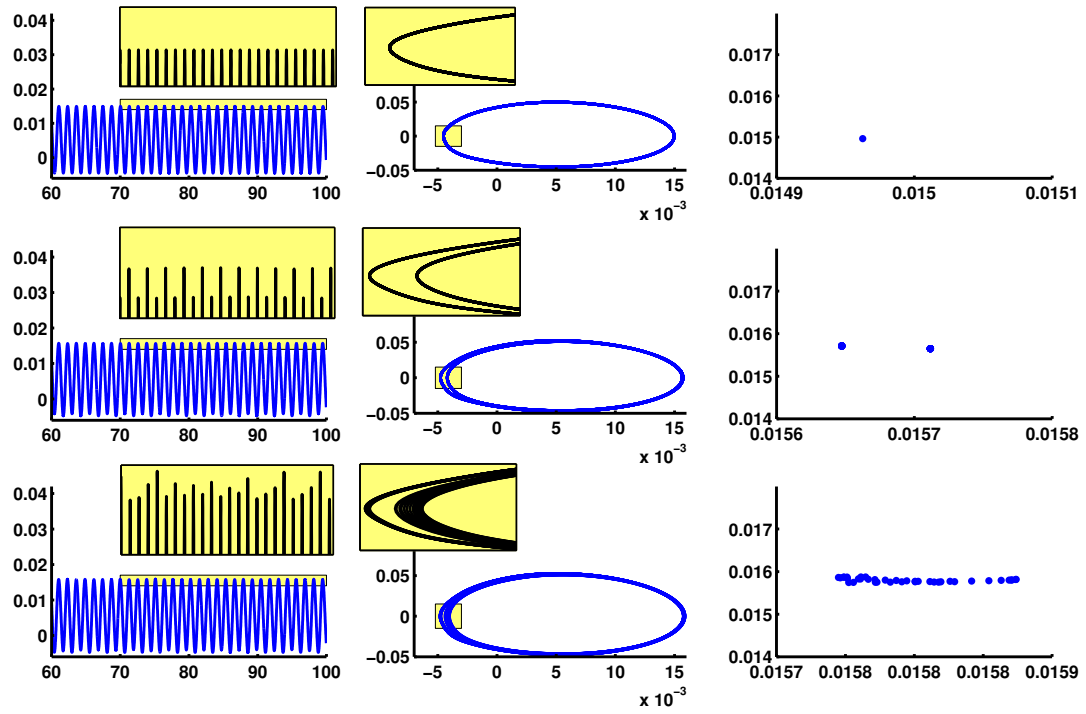


Figure 3.15: Attractors and Poincaré sections for oscillations on the large limit cycle (Fig. 3.10) confirm the appearance of subharmonics and chaos in the trapezoidal model. Each row shows the simulation results for increasing subglottal pressure (from top row to bottom, $P_s = (500, 530, 531.5)$ Pa, respectively) and initial values leading to the large limit cycle (namely $(x_{10}, v_{10}, x_{20}, v_{20}) = (-0.006, 0, 0, 0)$. All units are given in Table 3.2). Each column, from left to right, denotes: 1) the oscillation of the variable x_1 (and zoom on the plot around the maxima of oscillation), 2) the two-dimensional attractor obtained by plotting $v_1(t)$ versus $x_1(t)$ (and zoom around the minima of x_1) 3) modified Poincaré section. Transients were avoided by iterating the simulation on the same integration time $t = [0, 100]$ msec with the last values of the previous simulation as initial values. Modified Poincaré sections were obtained by plotting the pair $(x_{1j}^{max}, x_{1j+1}^{max})$, where x_{1j}^{max} represents the jth local maximum (obtained by means of parabolic interpolation) of the curve described by x_1.

3.4 Discussion

Our simulations show that rescaled biomechanical models originally developed to describe mammalian vocal fold vibrations can be adapted to model the bird syrinx. We assumed that both sound producing organs are excited by the same principle: in the opening phase a high pressure drives the vibrating structure apart and during closure the pressure is reduced due to the Bernoulli force. The fundamental frequency is governed by the mass and stiffness of the vibrating tissue.

Our simulations represent symmetric vibrations. Interestingly, the same model equa-

tions can be used to model a single vibrating structure (a "hemi-syrinx"). In this case only the sound intensity is reduced but onset pressure or intensity of harmonics are identical.

Control of harmonics in bird songs is widely debated (Klatt and Stefanski, 1974; Nowicki, 1987; Williams et al., 1989; Nowicki et al., 1992; Beckers et al., 2003a, 2004; Riede et al., 2004; Fletcher et al., 2004): some species have whistle-like songs and not much energy is found in the harmonics (Nowicki et al., 1992), whereas other species, such as the zebra finch (*Taeniopygia guttata*), display strong harmonics during song or calls.

In all two-mass model versions almost pure tones are found near the onset of vibrations (i.e. near the Hopf bifurcation line shown in Fig. 3.4). In the rescaled two-mass model, strong harmonics appear at higher pressures due to collisions. For a small upper mass and a rectangular geometry, collisions leading to strong harmonics can be avoided only near the phonation onset. At higher pressures counteracting forces would be required to diminish collisions. We hypothesize that the avoidance of strong collisions in song birds might be achieved by the medial tympaniform membranes (MTM) that are continuous with the inner vibrating labia (Fee, 2002). The role of the MTM has been widely speculated: it may play a distinct role especially for the generation of high-frequency sounds (Goller and Larsen, 2002), or contribute to the wave-like movements of the MVM (MTM +ML) observed in stroboscopic images of *in vitro* syringeal oscillations (Fee et al., 1998; Fee, 2002): as introduced in Section 2.1.2, wave-like movements result in a more efficient sound generation than possible with a one-mass oscillator (Titze, 1988, 1994; Fee, 2002).

In our model of the ring dove syrinx no collisions occur at default parameters. Consequently, harmonics are fairly weak. The smoother configuration and equal upper and lower masses counteract collisions even at relatively high pressures. This is presumably due to a stronger effect of the subsyringeal pressure acting on both masses. In a recent experimental paper, Riede et al. (2004) showed how varying suprasyringeal configurations can suppress the second harmonic in ring doves, and Fletcher et al. (2004) showed that the combined influence of trachea, glottis and upper esophagus acts as an effective band-pass filter that eliminates higher harmonics generated by the dove syrinx. Our simulations reveal that the configuration of the syrinx influences the intensity of overtones. Therefore, the amount of energy in the harmonics could also be controlled by syringeal muscles that directly affect the configuration of the syrinx (Goller and Larsen, 1997b; Elemans et al., 2004, 2006).

Around our default parameters given in Table 3.1 we found no voice instabilities. On the other hand, we have shown that our trapezoidal model exhibits coexistence of

attractors and jumps from harmonic rich spectra to more sinusoidal oscillations of the radiated sound pressure.

A previously published two-mass model of the songbird syrinx (Fee et al., 1998) demonstrated many instabilities such as period doublings, transitions from periodic to chaotic dynamics, as well as mode locking transitions. Unfortunately, we are unable to compare the behaviour of the two models in more details, because the governing equations and the parameter settings were not provided. Ring doves already exhibit stronger harmonics during inspiratory phonation compared to expiratory phonation even at low intensities (Gaunt et al., 1982; Beckers et al., 2003a; Elemans, 2004; Riede et al., 2004). This implies that asymmetries between outflow and inflow of the air have to be taken into account. In Appendix B, a detailed bifurcation analysis of the trapezoidal model will lead to a first attempt of simulation of the coo during inhalation, suggesting possible mechanisms involved in the generation of more rich dynamics. Instabilities such as frequency jumps are commonly observed in the ring dove coo (Beckers et al., 2003b; Elemans, 2004). Even our current model described in this paper exhibits already coexistence of different "registers". However, although intrinsic nonlinear properties of the syrinx add complexity to the level of motor control (Fee et al., 1998; Fee, 2002), only muscle control can explain the fast but gradual frequency modulations of ring dove coos (Elemans et al., 2004).

In this paper we have shown that the geometry and the rest position of the syrinx can influence the harmonic spectra drastically. Our simulations are a first step towards more realistic modelling of the syrinx. In the next chapter we will incorporate the dynamic control of associated superfast syringeal muscles. This will allow a quantitative comparison of observed bird songs and simulations.

4

Biomechanics and control of vocalization in a non-songbird

This chapter is an extended version of Elemans, C. P. H., Zaccarelli, R., and Herzel, H. (2007), "Biomechanics and control of vocalization in a non-songbird", J. Royal Soc. Interface (Elemans et al., 2007).

The neuromuscular control of vocalisation in birds requires complicated and precisely coordinated motor control of the vocal organ (i.e. the syrinx), the respiratory system and upper vocal tract. The biomechanics of the syrinx is very complex and not well understood. In this paper, we aim to unravel the contribution of different control parameters in the coo of the ring dove (*Streptopelia risoria*) at the syrinx level. We designed and implemented a quantitative biomechanical syrinx model that is driven by physiological control parameters and includes a muscle model. Our simple nonlinear model reproduces the coo, including the inspiratory note, with remarkable accuracy and suggests that harmonic content of song can be controlled by the geometry and rest position of the syrinx. Furthermore, by systematically switching off control parameters, we demonstrate how they affect amplitude and frequency modulation and we generate new experimentally testable hypotheses. Our model suggests that independent control of amplitude and frequency seems not possible with the simple syringeal morphology of the ring dove. We speculate that songbirds evolved a syrinx design that uncouples the control of different sound parameters and allows for independent control. This evolutionary key innovation provides an additional explanation for the rapid diversification and speciation of the songbirds.

4.1 Introduction

The neuromuscular control of vocalisation in birds requires complicated and well- orchestrated motor control of the vocal organ (i.e. the syrinx), the respiratory system and upper vocal tract (e.g., Wild, 1997). Especially songbirds, which learn their song from a tutor, provide an excellent and widely used model system for sensorimotor learning and human speech acquisition (Doupe and Kuhl, 1999). The biomechanics and neuromuscular control of the syrinx, which is a modification of the trachea and/or bronchi located at the bifurcation of the trachea into the bronchi, is very complex and not well understood. An emergent model organism to study the biomechanics of phonation and its control is the ring dove (*Streptopelia risoria*). The syrinx morphology of this species is relatively simple (Fig. 4.1) with only two paired muscles controlling its geometry, compared to 6-8 pairs in most songbirds (King, 1989). Because this non-songbird does not learn the syntax of its species-specific song as songbirds do (Nottebohm, 1972), it is of less interest to study vocal learning. However, juveniles of a closely related species still require motor practice to utter the final characteristic coo (Ballintijn and ten Cate, 1997), which indicates the individuals must go through numerous iterations to match their vocal output with some sort of neural template.

In pigeons, sound is produced in the syrinx by flow induced-oscillations of membranes (Goller and Larsen, 1997b; Larsen and Goller, 1999). The control of the fundamental frequency of the sound and the gating of sound elements happens at the level of the syrinx (Greenewalt, 1968; Goller and Suthers, 1996a,b). Despite the seeming simplicity of the dove coo, the syrinx demonstrates surprisingly complex dynamics, sound is produced during expiration and inspiration (Gaunt et al., 1982), the coo exhibits fast trills (Elemans et al., 2004) and frequency jumps (Beckers et al., 2003b).

Many forces act on the sound producing Lateral Tympaniform Membranes (LTMs) (Fig. 4.1b): pressure gradients (Gaunt et al., 1982; Beckers et al., 2003b) and syringeal muscles (Elemans et al., 2004, 2006) affect variations in geometry and tension (Gaunt, 1983). These forces affect the tension in the LTM (Gaunt, 1983; Fletcher, 1992), and most likely also the oscillation amplitude and phonation onsets. This complex interplay of forces is hard to dissect without a biomechanical quantitative approach.

Several conceptual models have been proposed to explain the biomechanics of coo production that use a separation of single physiological correlates with frequency modulation (FM) and amplitude modulation (AM) (e.g., doves: Beckers et al., 2003b; songbirds: Goller and Suthers, 1996a,b). This separation in functional attributes is most likely a too simple representation. Especially nonlinear oscillators, which are used to

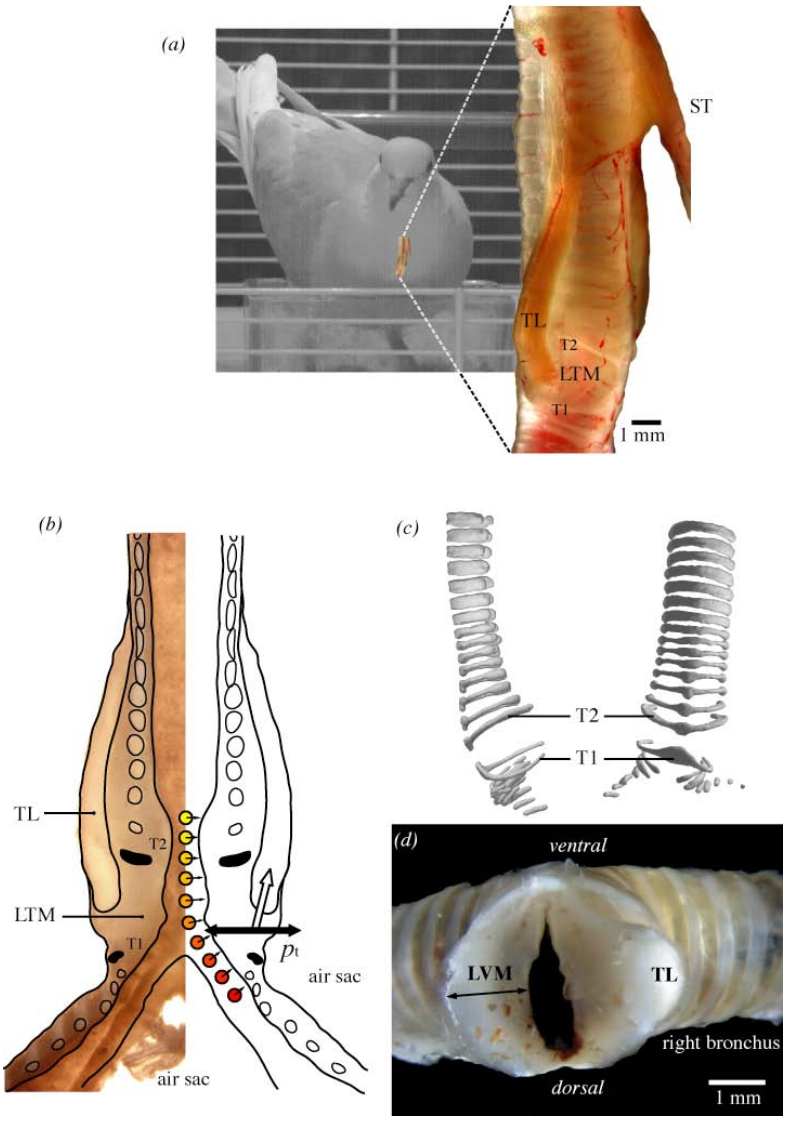


Figure 4.1: Syrinx morphology (a) toto view of ring dove syrinx. (b) longitudinal cross-section of ring dove syrinx. Several forces acting on the Lateral Vibratory Masses (LVMs): a pressure gradient from bronchus to trachea (circles), the transmural pressure P_t from air sac to syringeal lumen (black arrow), muscle activity (white arrow). (c) Micro-CT scan of syrinx. Note that due to its flattened shape, ring T2 is able to better withstand the torques generated by the m. tracheolateralis (TL). (d) Sagittal cross-section through a fresh syrinx slightly below the TL insertion. The right LVM is cut higher; some TL fibres are cut. The arrows indicate LVM width and length as presented in Table 4.1. LVM: Lateral Vibratory Mass, P_t: transmural pressure, T1-2: tracheal rings, TL: musculus tracheolateralis, ST: musculus sternotrachealis. The syrinx in Fig. 4.1a has been published in a different form as Online supplemental information accompanying Elemans et al. (2004) and in Elemans et al. (2006). The syrinx cross-section in Fig. 4.1b is modified from Zaccarelli et al. (2006).

model vocal fold oscillations (Ishizaka and Flanagan, 1972; Gardner et al., 2001; Fee et al., 1998; Laje and Mindlin, 2002; Mindlin and Laje, 2005; Laje et al., 2001, 2002), can show conversion from AM to FM in certain regions of their parameter space (e.g. the Duffing oscillator in Elemans et al., 2004).

In this paper, we aim to unravel the contribution of the different control parameters in the coo of the ring dove at the syrinx level. First, we briefly review the literature and aim to integrate and explain all observations into a novel biomechanical framework. Second, we characterize a standardized coo with all physiologically important parameters using earlier studies and additional measurements. Third, we implement a biomechanical syrinx model that is driven by the derived physiological control parameters. This model provides a consistency check for our formulated biomechanical framework. Furthermore, with our model, we can systematically switch on and off the control parameters to determine how they affect amplitude and frequency modulation. As such, we can also assess if it is reasonable to correlate amplitude and frequency modulation to separate physiological correlates and we can generate new experimentally testable hypotheses.

4.2 The biomechanics of phonation in ring doves

Over the last decades, the mechanics and physiology of phonation in doves have received some attention, but some data have been interpreted in an older framework missing recent insights (e.g. based on the assumption that other structures acted as the principal sound sources). Here, we aim to integrate all observations into a new framework.

Gaunt et al. (1982) show that abdominal muscles generate expiratory pressure pulses passing through the syrinx. However, during the coo, the beak and nares are closed and the air is collected in an inflating esophageal crop (Riede et al., 2004). Direct observations of the syrinx during induced phonation using brain stimuli in pigeons by (Goller and Larsen, 1997b) show that

- the Lateral Tympaniform Membranes (LTMs) are the sources of the sound and
- the syrinx is brought into a "phonatory position" by two paired syringeal muscles; the m. sternotrachealis (ST) and m. tracheolateralis (TL) as hypothesized earlier by (Gaunt et al., 1982).

Whereas the TL directly affects position in the LTM (Fig. 4.1a, b), the ST moves the entire syrinx downward (Goller and Larsen, 1997b). However, the LTMs in ring doves

are not thin membranes (see Section 4.5), but more resemble labia with larger masses. Therefore the denomination membrane seems not appropriate and we will argue that Lateral Vibratory Mass (LVM, Fig. 4.1) is a more appropriate name than Lateral Tympaniform Membrane. To avoid confusion, we will consistently use the abbreviation LVM from now on.

Although there are no direct observations in freely singing doves, several indirect measurements support the idea that the syringeal aperture is modulated and controlled by syringeal muscles. First, the duration of TL muscle activity is equal to the duration of sound elements during the trill (Elemans et al., 2004). Second, simultaneous pressure and flow recordings suggest that the syrinx is open during sound production and mostly closed in between sound elements: when sound starts, the bronchial pressure drops and when sound stops, flow decreases to almost zero (Gaunt et al., 1982; Beckers et al., 2003b). These observations can be explained by the hypothesis that TL contraction modulates the position of the LVM and as such the syringeal aperture (Elemans et al., 2004, 2006).

After the initial pressure drop due to opening of the syrinx, the bronchial pressure increases again during every sound element in the trill (Fig. 4.2a). Gaunt et al. (1982) suggest this is due to increased activity of abdominal muscles synchronized with the TL. Indeed the pressure slope starts increasing at the inflexion points on the negative slope of the pressure curve (Fig. 4.2a-c, local minima of $\frac{\mathrm{d}p}{\mathrm{d}t}$ in Fig. 4.2b), which can be indicative of a start of force generation by the abdominal muscles. We measured the delays between abdominal muscle EMG activity and pressure by digitizing the data presented in Gaunt et al. (1982). The delay between the onset of abdominal muscle EMG activity and the local minima in $\frac{\mathrm{d}p}{\mathrm{d}t}$ measures 7.7 ± 2.6 ms for the last five trills (Fig. 4.2). The pressure starts increasing 21.0 ± 7.4 ms after the onset of EMG activity. These delays are consistent with a typical delay between electrical activity (EMG) and mechanical muscle action for skeletal muscle (e.g., Norman and Komi, 1979; Komi, 2000; Carroll and Wainwright, 2006) and therefore the claim by Gaunt et al. (1982) is feasible. This again emphasizes that the generation of trills in doves demands a high level of synchronization of syringeal and respiratory muscle activation.

The parameters controlling the fundamental frequency are not well established. The common and most simple assumption is that the fundamental frequency ($\mathscr{F}_0$) of LVM oscillation is determined by tension in the LVM (Fletcher, 1992). Beckers et al. (2003b) report that the $\mathscr{F}_0$ correlates with the pressure in the air sac in which the syrinx is suspended; the interclavicular air sac (ICAS). Birds have a unique and elaborate air sac system for respiratory ventilation (Duncker, 1971). Pressure differences between air sacs have been reported during many activities such as normal respiration (Brackenbury, 1972), sound production (Brackenbury, 1972) and locomotion (Boggs et al., 1998, 2001). Beckers et al. (2003b) hypothesize that ICAS pressure modulates the ten-

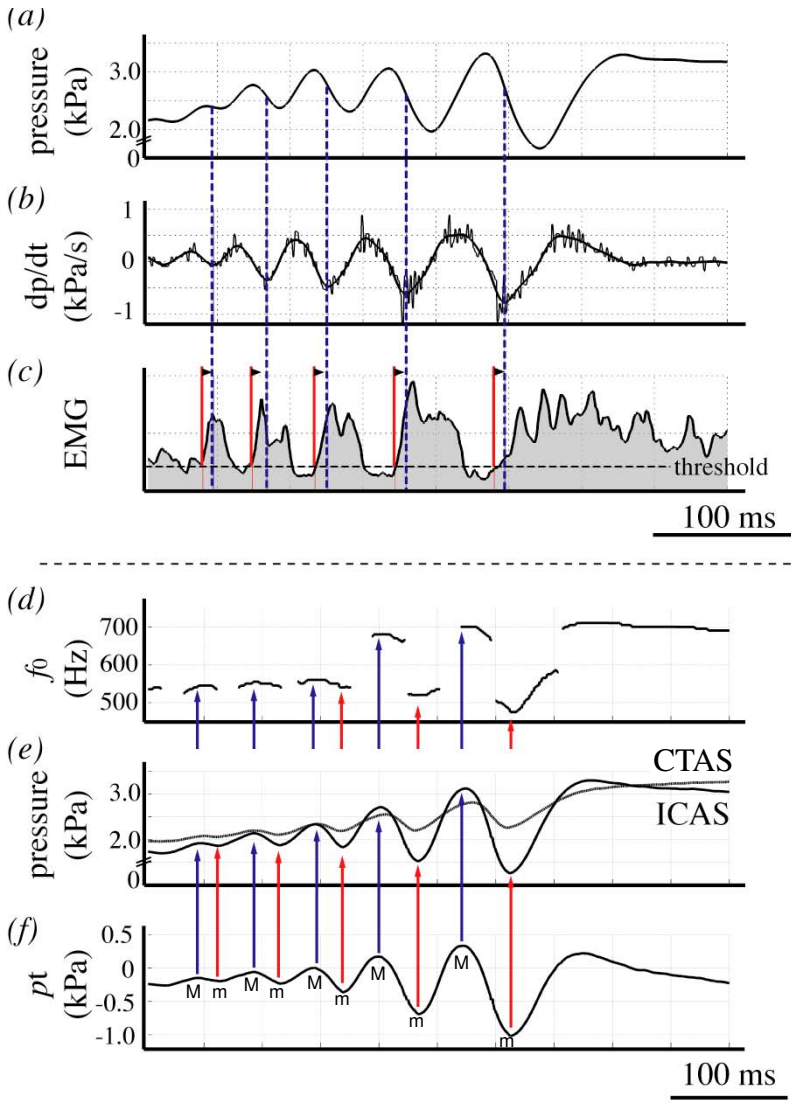


Figure 4.2: Function of abdominal muscles and air sac pressure derived from literature. (a)-(c) Data digitized from Gaunt et al. (1982). Abdominal muscles contract to increase bronchial pressure. (a) posterior thoracic air sac (PTAS) pressure, (b) time derivative of air sac pressure in (a) and (c) EMG activity of abdominal muscles (integrated trace, time constant 1 ms). Onset of muscle activity (red solid lines) commences 7.7 ± 2.6 ms (mean ± 739 S.D., n=5) before the inflexion points of air sac pressure (blue dashed lines corresponding to local minima in (b)). (d)-(f) Data digitized from Beckers et al. (2003b). Transmural pressure waveform is in phase with interclavicular air sac (ICAS) pressure waveform. (d) Fundamental frequency ($\mathscr{F}_0$) during trill. (e) ICAS (solid line) and cranial thoracic air sac (CTAS) pressure (dotted line) patterns. (f) calculated transmural pressure (P_t) using eq. 4.1 and the data in (e). Local maxima (blue arrows, corresponding to the label "M") and minima (red arrows, corresponding to the label "m") for P_t, ICAS and fundamental frequency coincide.

sion in the LVM. However, tension in the membrane is the result of the net force that acts on the membrane. When looking at the forces acting on the LVM (Fig. 4.1b), it becomes clear that any pressure differences between the air sac surrounding the syrinx (the interclavicular air sac) and syringeal lumen causes a net force to act on the LVM. This so-called transmural pressure P_t affects the tension in the membrane (Bertram and Pedley, 1982; Bertram, 2004). We define the transmural pressure as external pressure minus upstream pressure, where we take ICAS pressure as external pressure and assume that caudal thoracic air sac (CTAS) equals upstream or bronchial pressure (see Section 4.3.1):

$$P_t = P_{\mathrm{ICAS}} - P_{\mathrm{CTAS}}, \tag{4.1}$$

where P_{ICAS} and P_{CTAS} are the pressures in ICAS and CTAS respectively. When we derive P_t from the data presented by Beckers et al. (2003b), we see that the calculated P_t is in phase with the pressure in the ICAS (P_{ICAS}) (Fig. 4.2d-f). Therefore correlation techniques cannot differentiate between P_{ICAS} or P_t as modulators and as such the function "frequency modulator" cannot be attributed to either one of them using correlation. Considering the mechanics, the LVM tension is affected by both 1) the transmural stress caused by a pressure differential between the bronchus and ICAS and 2) the stress exerted by muscles.

To summarize, we hypothesize that

- the tracheolateralis muscle (TL) modulates the position of the lateral vibratory mass (LVM) and as such the syringeal aperture and
- both the transmural pressure and TL stress affect tension in the LVM.

4.3 Compilation of a standardized coo

We compiled a quantitative dataset of all the parameters affecting sound production at the syrinx level. Many doves have 3 different coo types; the nest-, bow- and perch coo (Goodwin, 1983). We did not differentiate between these coos because in ring doves there are no apparent differences in overall acoustic structure (Beckers et al., 2003b). We combined published and additionally collected physiological and morphological data to construct a "standard coo" (Fig. 4.3). Our primary focus is the trill where fastest and most complex interactions occur. If we look at the forces acting on the syrinx membranes (Fig. 4.1b), the most important physiological control parameters are 1) the bronchial-tracheal pressure gradient, 2) the transmural pressure difference and 3) the

stress exerted by syringeal muscles. It should be noted that all these control parameters are the ultimate result of muscle action and therefore all under direct neural control. All pressures are noted as gauge pressures, i.e. pressure relative to atmospheric pressure. To construct all synchronized time dependent parameters of the standard coo, we start with simultaneously recorded signals of emitted sound, caudal thoracic air sac (CTAS) pressure and electromyograms (EMG) of the syringeal muscles. These signals were recorded as part of previous studies (Elemans et al., 2004, 2006).

4.3.1 The bronchial-tracheal pressure gradient

To understand the vibrations of the Lateral Vibratory Membrane (LVM), we need to know upstream (i.e. bronchial) and downstream (i.e. tracheal) pressure because the pressure drop along in the LVM defines the oscillation regime, analogous to collapsible tube systems (e.g., Bertram, 2004; Grotberg and Jensen, 2004) and human vocal fold models (Titze, 2002).

To estimate upstream pressure, i.e. the bronchial pressure, the common assumption is that pressure in two easily accessible air sacs, either the pressure in the posterior thoracic air sac (PTAS) or caudal thoracic air sac (CTAS), equals bronchial pressure for most species of birds (e.g., Gaunt et al., 1982; Suthers et al., 1994; Goller and Suthers, 1996a; Mindlin et al., 2003; Suthers and Zollinger, 2004). To our knowledge, no record of direct bronchial pressure exists of any bird species during song. We studied silicone casts of the air sac system in ring doves, and found that the connections or ostia of the lung with CTAS are close to the bronchi (1-3 mm) cf. Duncker (1971). Therefore also in this case, it is safe to assume that the pressure in the CTAS is a good estimate for bronchial or subsyringeal pressure. Published PTAS and CTAS pressure patterns for ring doves are highly similar (compare PTAS pressure in figures 5, 6 and 7 in Gaunt et al., 1982 with CTAS pressure in figures 1 and 2 in Beckers et al., 2003b and CTAS pressure in Fig. 4.3b). A small phase shift between the pressure waveforms of air sacs cannot be ruled out, but from what is known from morphology and airflow pattern models there is no reason to assume large differences in amplitude (Brackenbury, 1972).

The maximal pressure in the PTAS and CTAS is 3.5 kPa (=35 cmH_2O), with a slight variation introduced by the emotional state of the individual bird (Gaunt et al., 1982). Therefore, it is reasonable to set the maximal pressure for CTAS, PTAS and ICAS in the air sac system to 3.5 kPa. To determine the bronchial-tracheal gradient, we measured CTAS pressure by canulating CTAS in six male doves (data collected in collaboration with Dr. F. Goller as part of previous studies: Elemans et al., 2004). Our results are consistent with earlier measurements; maximal pressure is 3.5 kPa and the temporal pattern is highly similar (figure 3b).

Tracheal pressure gradually increases during the coo, because the air is not exhaled

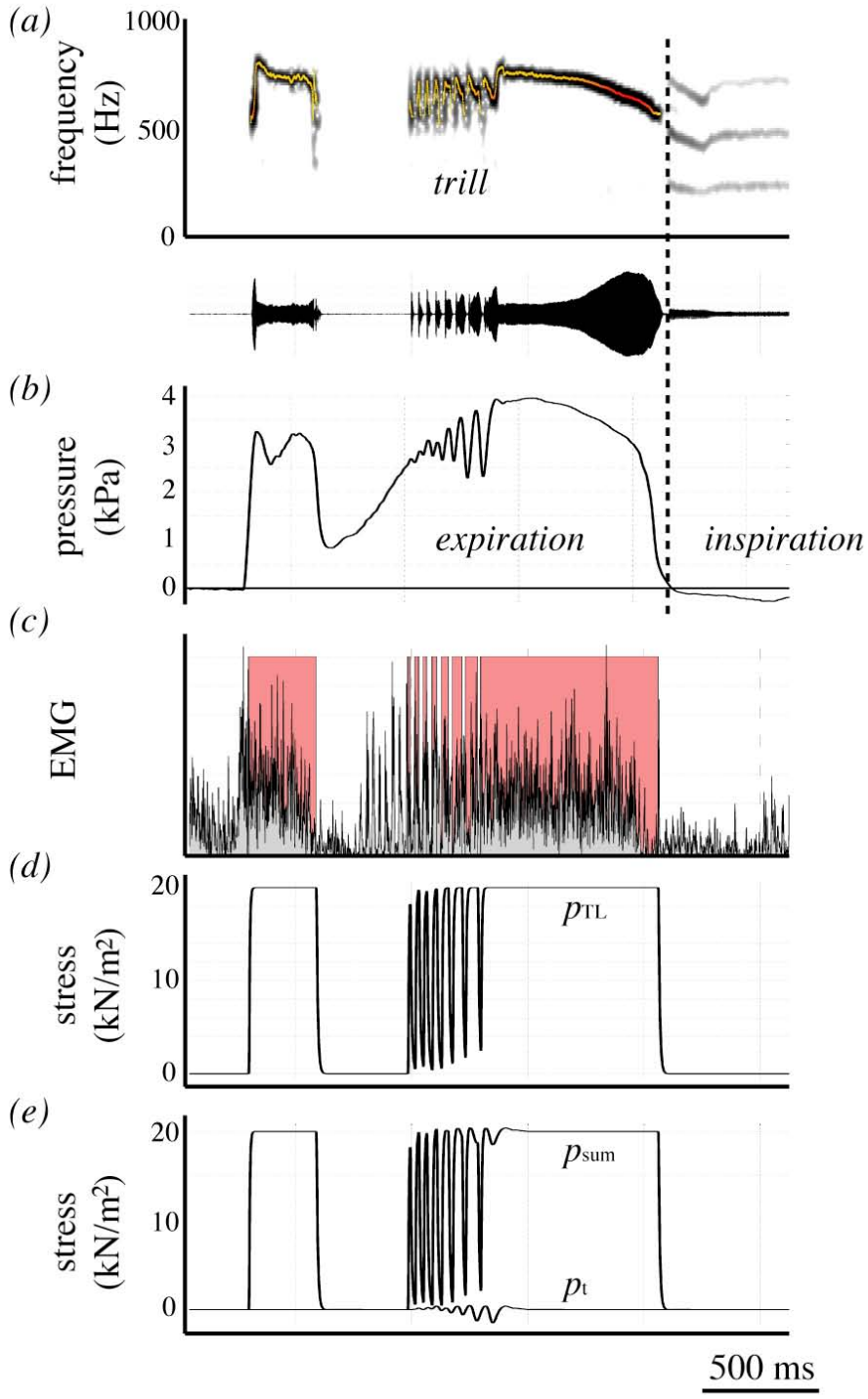


Figure 4.3: Standard coo (a) spectrogram and oscillogram of dove coo. Superimposed on the spectrograms is a fundamental frequency analysis (light line). (b) bronchial pressure. The vertical dotted line indicated the transition from expiration to inspiration. (c) rectified EMG of tracheolateralis muscle (TL). The shaded areas in background (red areas) represent the TL activity associated with sound elements. (d) stress generated by TL (e) transmural stress (bottom trace) and sum of transmural and TL stress. P_t: transmural pressure, P_{TL}: stress generated by TL, P_{sum}: sum of P_t and P_{TL}.

but collected in the esophagus (Gaunt et al., 1982; Riede et al., 2004). The peak pressure for tracheal pressure rarely exceeds 0.5 kPa (=5 cmH_2O) and often only slightly exceeds the respiratory exhalation pressure of 0.1 kPa (=1 cmH_2O). Because this pressure is low compared to the bronchial pressure, we will assume that the downstream pressure equals atmospheric (0 kPa). This simplification implies that the CTAS pressure defines the bronchial-tracheal pressure gradient.

4.3.2 The transmural pressure differential

As discussed above, any pressure difference between the bronchial-tracheal gradient on the one hand and the pressure in the ICAS on the other, results in a net force on the LVM and therefore in a tension change in the membrane. As can be seen in Fig. 4.2e and 4.2f, the transmural pressure (P_t) oscillates in phase with P_{ICAS} and almost in phase with P_{CTAS}. To reconstruct P_t from P_{CTAS}, we retain the oscillation in P_{CTAS}, and normalized the magnitude to 1 kPa as measured in Fig. 4.2f (bottom trace in Fig. 4.3e).

4.3.3 Syringeal muscles

In ring doves, two syringeal muscles affect LVM position; the paired m. tracheolateralis (TL) and m. sternotrachealis (ST) (Fig. 4.1). Although the ST seems important for stability of the syrinx (crows: Chamberlain et al., 1968; thrashers: Goller and Suthers, 1996a,b; doves: Elemans et al., 2006), they do not act directly on the LVM and we will ignore them here to simplify matters. We use the following algorithm to construct the stress exerted by the TL. First, we define blocks of muscle on/off activity by creating a binary on/off signal from the measured EMG signal (Fig. 4.3c). The threshold needed to calculate the binary signal was set to the mean value of the noise level plus three times the standard deviation. Second, we fit exponential rise and fall curves through the tetanic isometric contraction data of Elemans et al. (2004) and calculate their time constants. The maximal stress is set to 20 kN/m^2 (= maximal isometric stress from Elemans et al., 2004). Third, we numerically construct the stress signal using the extracted time constants and the on/off EMG signal (Fig. 4.3d). We omit the pre-trill pulses of muscle activation (Fig. 4.3c), and include only pulses associated with sound elements. The total stress acting on the LVM is $P_{sum} = P_t + P_{\mathrm{TL}}$ (Fig. 4.3e).

4.3.4 Morphology

The thickness of the LVM varies along the syrinx. We assume that action of TL mostly affects the mass between tracheal rings T1 and T2 (Fig. 4.1b), because the tracheal rings T2 - T20 form a stiff box and both T1 and T2 are flattened rostro-caudally (Fig. 4.1c). Furthermore, endoscopic observations by Goller and Larsen (1997b) show that in the

rock pigeon (*Columbia viva*), action of TL and ST especially affects the position and tension of the LVM between T1 and T2. We measured the dimensions and estimated mass of the LVM in four adult male ring doves. LVM height is measured between T1 and T2 (Fig. 4.1b), and LVM length and width are measured slightly below the TL insertion between tracheal ring T1 and T2 (Fig. 4.1d) using a MicroPhot-FXA microscope (NIKON) and ANALYSISpro software.

4.4 Biomechanical model description

We modeled one unilateral LVM with the trapezoidal model described in the previous chapter. Because two-mass models have been studied intensively for human and nonhuman phonation, they provide a well-established starting point. Now we contrast our approach to two previously developed types of syrinx models introduced in Section 2.3.3 (for a review of syrinx models see Elemans et al., 2003). The first mathematical model of the syrinx by Fletcher (1988) models the LVM as a thin membrane. However, the LVMs in ring doves are not thin membranes (see results in Section 4.5). Therefore the denomination membrane seems not appropriate and the structures more closely resemble labia with larger masses (Fig. 4.1 and Table 4.1). Furthermore, as also noted by Laje et al. (2002) (see Section 2.3.3), in order to produce vocalizations with a natural spectral content, the driving pressure in Fletcher's model is composed of a harmonically complex pressure waveform, which does not correspond to the simple pressure traces. Therefore, the model of Fletcher (1988) is not a good starting point to model the ring dove syrinx. The second group consists of elegant single-mass models developed by Mindlin, Laje and co-workers (see Section 2.3.3). These models are deliberate simplified to the bare minimum to study nonlinear effects caused by control parameters (Gardner et al., 2001; Laje et al., 2001, 2002; Mindlin et al., 2003; Mindlin and Laje, 2005). However, the control parameters cannot be easily translated back to physical properties.

Fig. 4.4 shows cartoons of our biomechanical syrinx model in three different configurations; convergent, rectangular and divergent. The subsyringeal pressure P_s generates a pressure force on the plates and pushes the masses outward, loading the two springs and dampers. The two masses are interconnected by a single spring with coupling constant k_c. Above a threshold value, the pressure induces self-sustained oscillations of the masses. The opening and closing of the syringeal aperture creates pressure waves i.e. sound.

Description	Value (mean ± S.D.)	Corresponding model parameter (Fig. 4.4a)
height T1-T2 (mm)	2.2 ± 0.4 ($n = 4$)	$d_1 + d_2 + d_3$
maximal width of LVM between T1-T2 (mm)	1.6 ± 0.3 ($n = 8$)	
depth (mm)	3.2	ℓ
tracheal diameter (mm)	3.0	$2w$

Table 4.1: Geometry of the LVM. For measurements see Fig. 4.1b and 4.1d.

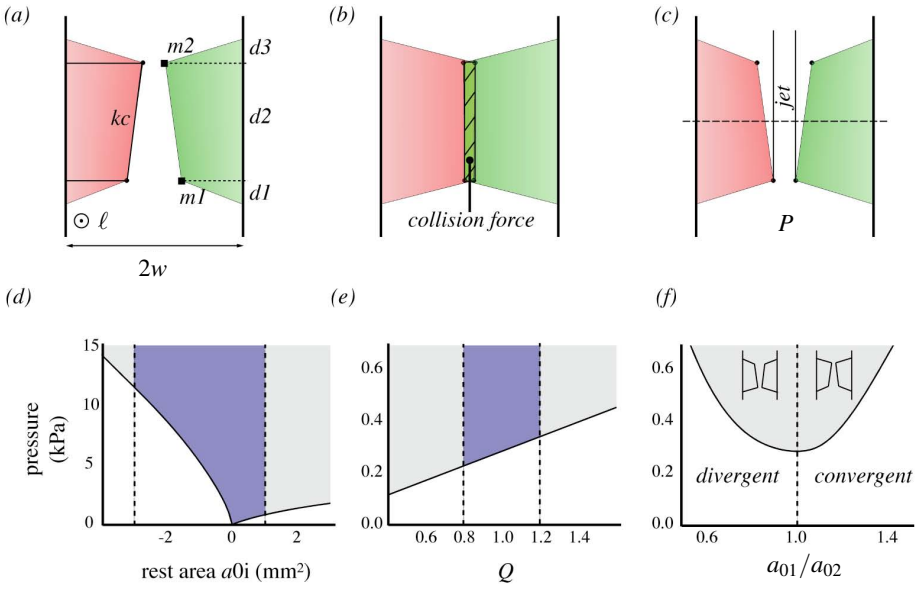


Figure 4.4: Biomechanical syrinx model. The biomechanical syrinx model in convergent (a), rectangular (b) and divergent (c) shape. Horizontal line in (c) is the midline of the model. Parameters are explained in text and listed in Table 4.1 and 4.2. (d)-(f) Bifurcation diagrams of oscillation onset pressure versus (d) rest area, (e) tension parameter Q and (f) shape of the rest configuration (or prephonatory shape). The shaded region above the bifurcation line represents self-sustained oscillation. The dark area in between the dotted lines (blue area) indicates the parameter range used in the simulations (see Appendix C).

Following classic two-mass models for human sound production, the pressure forces within the syringeal lumen that act directly on the labia are obtained using Bernoulli's law and the jet separation assumption (see Section 2.1.2 and Chapter 3). We summarize here in few points the basic assumptions described extensively in Chapter 3:

- We assume that the air stream separates from the vocal fold surface at the smallest syringeal opening area to form a jet (Pelorson et al., 1994; Titze, 1994). This jet keeps the pressure above the jet separation point close to zero. For the divergent shape, a non-fixed separation point improves the simulated sound (Pelorson et al., 1994) and this method is used in a study of vocal fold model with applications to

prostheses design (Lous et al., 1998). However, because we have no information on the flow field inside the syrinx, we assume a fixed separation point in the divergent case cf. previous models (Ishizaka and Flanagan, 1972; Story and Titze, 1995; Steinecke and Herzel, 1995).

- Our syrinx model deviates from the classical two-mass models in two important points.

  ① we assume symmetry between the masses ($m_1 = m_2$, $k_1 = k_2$, $r_1 = r_2$).

  ② we introduce a more smooth geometry; on each unilateral side of the LVM the two masses are linked with three mass-less plates (Fig. 4.4a-c).

- We define a collision force that pushes the two contralateral LVMs away when they collide. The collision force is a linear function of the area which the two LVMs overlap (Fig. 4.4b).

- To evaluate force on each mass, we assume that all forces acting below the syrinx midline (horizontal line in Fig. 4.4c) act on the lower mass m_1, whereas all the forces acting above act on the upper mass m_2 (a more detailed description of the equations of motion and pressure evaluation of our model can be found in Appendix A).

- The flow derivative $\dot{U} = \frac{\mathrm{d}U}{\mathrm{d}t}$ is used as an approximation of the suprasyringeal sound (Ishizaka and Flanagan, 1972).

4.4.1 Implementation of time dependent parameters

In Section 4.2, we argue that the syringeal muscles, bronchial pressure and the transmural pressure affect the geometry and tension of the LVM. We hypothesize that

① activity by the TL muscle modulates syringeal aperture and

② the sum of the transmural pressure and muscle stress ($P_{sum} = P_t + P_{\mathrm{TL}}$) modulates LVM tension.

We implemented the above hypotheses in our model, as described in detail in Appendix C. The modulation of rest areas (i.e. syringeal aperture) is a linear function of force exerted by the TL (Fig. 4.3d). We implemented tension modulation of the LVM by means of parameter $Q(t)$, which re-defines the masses and the stiffness of the LVM as a function of time: $m(t) = \frac{m}{Q(t)}$, $k(t) = kQ(t)$, where mass m and stiffness k are the default values (Table 4.2). Several implementations of Q have been used in Ishizaka and Flanagan (1972) and Steinecke and Herzel (1995). The definition of Q is motivated

Description	Symbol	Value
width of syringeal base (mm)	w	1.5
depth of syringeal base (mm)	ℓ	3.0
height of mass 1 (mm)	d_1	0.4
height of mass 2 (mm)	$d_1 + d_2$	2.4
LVM height (mm)	$d_1 + d_2 + d_3$	2.8
mass* (g)	m $(m_1 = m_2)$	$1.7 \cdot 10^{-3}$
stiffness* ($\frac{g}{ms^2}$)	k	$22.0 \cdot 10^{-3}$
damping constant ($\frac{g}{ms}$)	r	$1.2 \cdot 10^{-3}$
coupling constant* ($\frac{g}{ms^2}$)	k_c	$6.0 \cdot 10^{-3}$
rest area of mass 1 (mm^2)	a_{01}	0.3
rest area of mass 2 (mm^2)	a_{02}	0.3

*Parameters are rescaled during coo according to eq. 4.2 (for details, see Appendix C)

Table 4.2: Model parameters, for geometry parameters see Fig. 4.4, Chapter 3 and Appendix A.

by the relation $\mathscr{F}_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$, which defines the natural oscillation frequency ($\mathscr{F}_0$) of a single mass-spring system. For a detailed analysis of natural frequencies of oscillations of the classical two-mass model see Titze (1976, 1994). The $\mathscr{F}_0$ as a function of time is

$$\mathscr{F}(t) = \frac{1}{2\pi}\sqrt{\frac{kQ(t)}{\frac{m}{Q(t)}}} = Q(t)\mathscr{F}_0 \tag{4.2}$$

We systematically explored the bifurcation behaviour of our model (for details, see Appendix B) using the XPPAUTO software package (available at `www.math.pitt.edu/~bard/xpp/xpp.html`).

4.4.2 Scaling functions

As discussed above, the measured values of the stress due to TL are in the range [0,20] kN/m^2. Simulations of the model indicate that the corresponding range at rest areas a_{0i} is between -3 and 1 mm^2. We introduce a scaling factor to transform the range of TL values linearly to the range of rest areas. In a similar way we rescale the range of P_{sum} to Q values between 0.8 and 1.2 to match the frequency range of the coo (see Fig. 4.4e). A detailed description of the scaling functions $a_{0i}(t)$ and $Q(t)$ can be found in Appendix C. We assume that $F(t)$ varies tissue mass and stiffness around their default values. This requires a slight increase of the default m and k values, closer to the measured values with respect to the set of parameters in Zaccarelli et al. (2006). The damping ratio ($\zeta = \frac{r}{2\sqrt{km}}$) - a measure for viscous loss of energy in the focal fold tissue - was kept equal to the values set for human vocal cords (Ishizaka and Flanagan, 1972; Steinecke

and Herzel, 1995). Because of the increased mass and stiffness values, we slightly increased the damping r accordingly (see Table 4.2).

4.5 Model results

4.5.1 Model performance

We hypothesized that the tracheolateralis muscle (TL) modulates the syrinx aperture and that the net stress working on the LVMs modulates their tension. Table 4.1 lists the morphological input parameters. The LVM has a thickness (width) of 1.6 ± 0.3 mm (n=8). Therefore, the term "membrane" in the previously used term "Lateral Tympaniform Membrane" seems not appropriate regarding the connotation in both physics and histology. As introduced before, we will therefore use the abbreviation Lateral Vibratory Mass (LVM). We explore the model behaviour during expiration by studying free parameters not provided by measurements (for details, see Appendix B). Such parameters allow us to maintain the necessary model assumptions as well as to explore possible different model behaviours. These parameters are rest areas a_{0i}, and shape of the rest configuration ($\frac{a_{01}}{a_{02}}$). Bifurcation diagrams of these parameters and tension Q regarding the onset pressure can be seen in Fig. 4.4d-f. The shaded area above the bifurcation line represents oscillations of the LVM. These diagrams reveal that:

- onset of oscillations decreases when the rest area approaches zero (see Fig. 4.4d). Therefore oscillation is possible at low pressures when the LVMs are adducted into the lumen.

- For rest areas close to zero, a slightly divergent pre-phonatory shape ($\frac{a_{01}}{a_{02}} < 1$, Fig. 4.4f) leads to a lower onset of oscillations. Furthermore, a slightly divergent pre-phonatory shape leads to more instabilities, such as period doublings, at low pressure values (see Appendix B).

Fig. 4.5 shows the whole coo synthesized with our biomechanical model using the "standard coo" input signals from Fig. 4.3. Both the sound oscillogram and spectrogram closely resemble the sound uttered by the dove (Fig. 4.6). When we look closer at the trill, the most dynamic part of the coo, our model reproduces the up and down modulation during on-, and offsets during trill elements and the gradual modulation in between very well (Fig. 4.5e). Our model also reproduces the sharp onset of trill elements and amplitude modulations in the time domain (Fig. 4.5f). Furthermore, flow

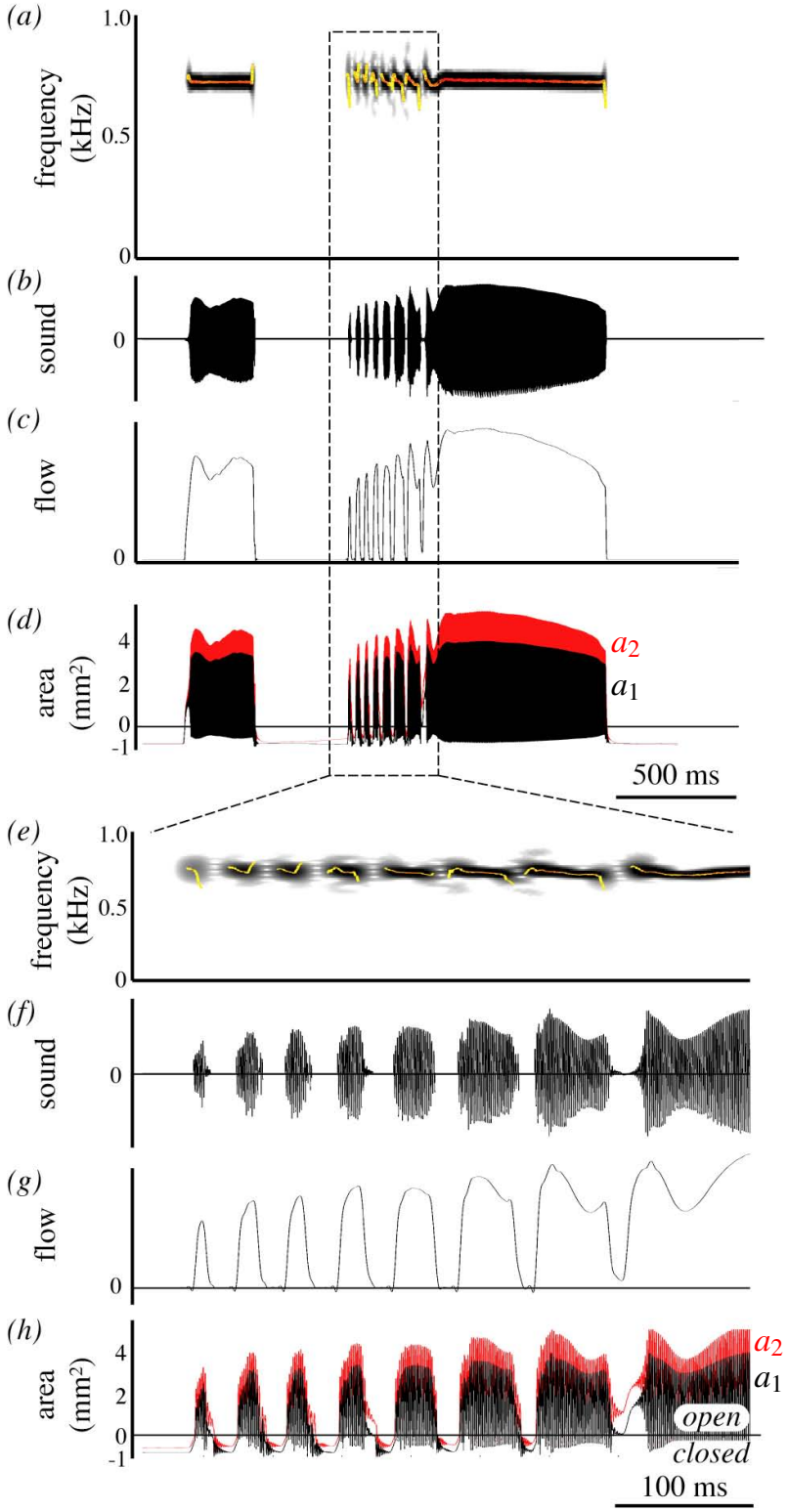


Figure 4.5: Simulation of the ring dove coo (a) sound spectrogram, (b) sound oscillogram, (c) flow in the trachea, (d) area of syringeal aperture for mass 1 (a_1) and mass 2 (a_2). (e)-(h) same signals zoomed in on the trill. Superimposed on the spectrograms is a fundamental frequency analysis. Negative areas in (d) and (h) indicate that the syrinx is closed in between sound elements.

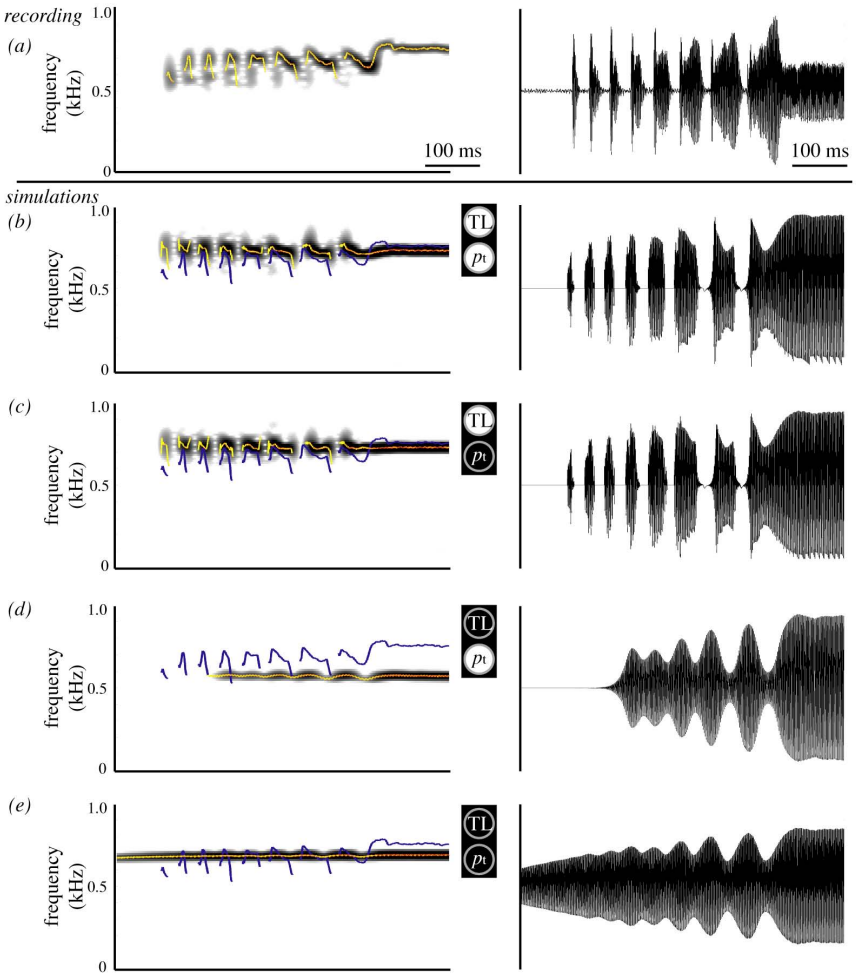


Figure 4.6: The effect of control parameters on the trill (a) Recorded sound in vivo. (b)-(e) model simulations with (c) no transmural pressure (d) no TL activity, and (e) no transmural pressure and no TL activity. In (c), the rest areas are constant ($a_{0i} = 0.3$ mm^2, see Table 3.2 on p. 54). In (d), $Q(t) = 1$ (see eq. 4.2) and rest areas are constant ($a_{0i} = 0.3$ mm^2). Superimposed on spectrograms are: dark blue line, $\mathscr{F}_0$ analysis of recorded coo in (a); light line, $\mathscr{F}_0$ analysis of current signal. Left panels: sound spectrograms, Right panels: sound oscillograms.

is zero (Fig. 4.5g) and the syringeal aperture areas are negative in between trill elements (Fig. 4.5h), implying that the syrinx is closed.

4.5.2 In silico experiments with control parameters

With our model, we can systematically switch on and off the control parameters transmural pressure P_t and TL activity to determine how they affect amplitude modulation (AM) and frequency modulation (FM). Fig. 4.6 shows what happens when we deactivate the TL muscle, the transmural pressure modulation and both. First, we observe that turning off the transmural pressure, has a very small effect on either the onset of sound elements, AM or FM (compare Fig. 4.6b and 4.6c). Also during the fast trills,

both AM and FM closely resemble the original coo recording in Fig. 4.6a. However, switching off the TL has profound effects on both gating and FM (Fig. 4.6d and 4.6e), compared to the original recording. Obviously there is no distinct gating of sound elements anymore. However, the overall amplitude of the sound is still modulated, but now by the bronchial pressure. The remaining small depth of FM is the result of changes in transmural pressure. By turning off both the TL and transmural pressure as modulators of Q (Fig. 4.6e), practically no FM remains and the amplitude is modulated by the bronchial pressure.

4.5.3 Asymmetry between expiration and inspiration

In a typical ring dove coo, nonlinear transitions such as period doublings and chaos are rare during expiration, and the radiated output signal is almost purely tonal. The amplitude of the harmonics $2\mathscr{F}_0$ and $3\mathscr{F}_0$ during a steady part of the second syllable for the coo depicted in Fig. 4.3a, measure -55 and -40 dB respectively compared to $\mathscr{F}_0$. These values are very similar to published values: $2\mathscr{F}_0$ and $3\mathscr{F}_0$ are both 40 dB lower in amplitude respectively in Riede et al. (2004) and 40 and 45 dB lower in Beckers et al. (2003a). The inspiratory note following expiration on the other hand, is characterized by richer harmonic spectra (Fig. 4.3a). The amplitude of the harmonics $2\mathscr{F}_0$ and $3\mathscr{F}_0$ are 8 and 12 dB higher compared to the fundamental frequency.

Our model also exhibits more rich harmonic spectra during inspiration than during expiration (Fig. 4.7, see Appendix B for details). The first and second harmonic measure respectively -22 dB and -34 dB in amplitude compared to the fundamental during expiration, and -6 dB and -12 dB during inspiration. These results imply that the geometry of the model can influence the harmonic sound spectra, because the shape of the LVM differs between simulated expiration and inspiration.

4.6 Discussion

We establish a novel biomechanical framework for sound production in ring doves (*Streptopelia risoria*). Integrating published experimental data, we hypothesize that

① the tracheolateralis muscle (TL) modulates the position of the Lateral Vibratory Masses (LVM) and as such the syringeal aperture and

② both TL stress and the transmural pressure affect tension in the LVM.

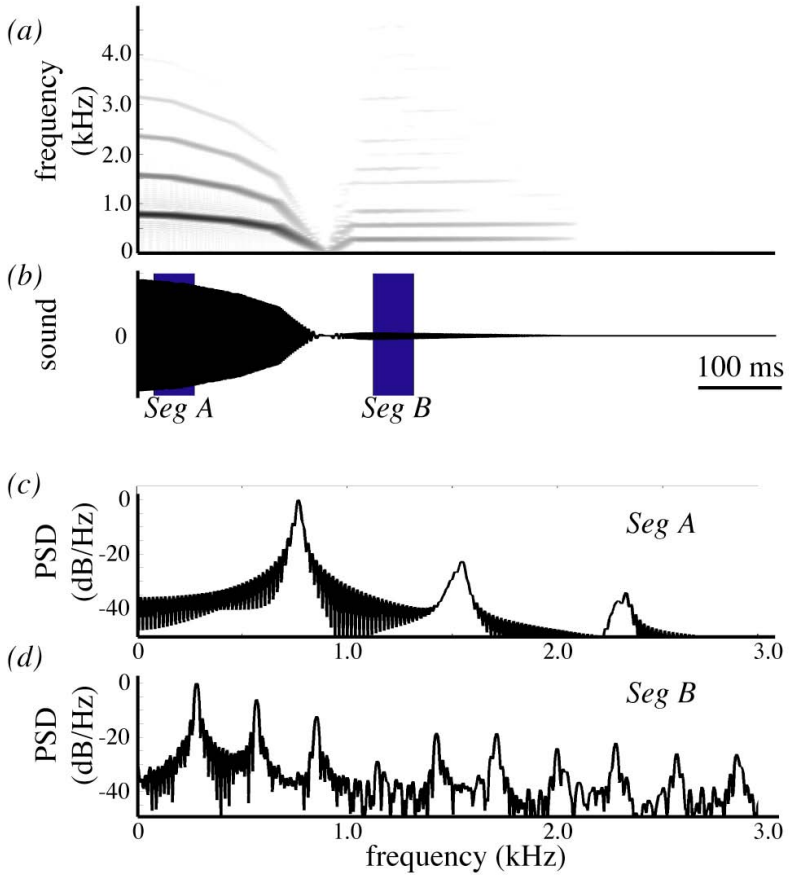


Figure 4.7: Simulation of the asymmetry expiration / inspiration: (a) Spectrogram and (b) oscillogram of inspiratory note simulation. (c) Power spectral density estimations of segment A (Seg A) during expiration and (d) segment B (Seg B) during inspiration. Segments are taken from (b) as indicated.

We designed and implemented a quantitative biomechanical model to confirm the validity of these hypotheses. Our simple nonlinear model reproduces the coo with great accuracy and as such provides strong evidence for our biomechanical framework. This is not a trivial result, because nonlinear dynamical systems make the principle of superposition no longer applicable: different excitations may result in different time-courses, described by nonlinear dynamics theory (Arrowsmith and Place, 1990; Strogatz, 1994; Titze et al., 1993). As described in Chapter 3, in simulations with time-constant parameters (with the previous version of our model), we showed that different parameter values can lead to distinct oscillation patterns, and small perturbations induced jumps from one attractor to the other (Zaccarelli et al., 2006). In principal, the interplay of forces that affect rest areas, stiffness and subsyringeal pressure can lead to nonlinear phenomena such as subharmonics, biphonation or deterministic chaos (Fee et al., 1998; Wilden et al., 1998; Fitch et al., 2002). In a typical ring dove coo, register transitions to aphonia, period doublings or chaos are rare during expiration or inspiration. The

simulations with physiological input parameters (Fig. 4.5 and 4.6) do also not exhibit transitions to different dynamical regimes. Ring dove coos can also exhibit punctuated frequency jumps (Beckers et al., 2003b), but these are not reproduced by our model.

We present the first model of bird song that implements a muscle model to calculate force produced by the muscle. The force production by the muscle is based on its actual contractile properties and activation and not simply a function of EMG activity.

The thickness (width) of the LVM is 2-16 fold higher than reported in other species. In the closely related rock pigeon (*Columbia viva*), the LVMs are masses of connective tissue measuring up to 0.1 mm thick (Goller and Larsen, 1997b). In ducks, reported values range from 360 μm (Frank et al., 2006) to 750 μm (Warner, 1972). The LVMs in ring doves seem to more closely resemble the consistency and geometry of the mammalian vocal folds or songbird's labia. Until direct observations were made in situ of the oscillating LVMs by (Goller and Larsen, 1997b), the Medial Tympaniform Membranes (MTM) were thought to be the vibrating sound sources (see Section 2.3). These much thinner membranes (5-20 μm; Casey and Gaunt, 1985) inspired the implementation of quantitative models where the sources were modelled as a membrane (Fletcher, 1988) or string (Casey and Gaunt, 1985), as described in Section 2.3.3. Our model provides a better description of the actual syrinx geometry.

Larsen and Goller (1999) observed 0.5 - 1.0 mm oscillations of the LVM in pigeons. We can derive the maximal oscillation amplitude of one modelled LVM from Fig. 4.5d and the relation $a_2 = 2\ell x_2 + a_{02}$ (see Appendix A). We have to consider mass 2, because Larsen and Goller (1999) had a top-view of the syrinx. The maximal oscillation amplitude of our modelled LVM is 0.8 mm ($\frac{4.8}{2 \cdot 3.0}$), which corresponds very well with the observed 0.5 - 1.0 mm. Larsen and Goller (1999) also suggest that a smaller amplitude vibration (0.001 - 0.1 mm) exist on top of the 0.5 - 1.0 mm oscillation using a vibration detector that detected light reflectance of the LVM. This signal is hard to interpret in terms of lateral movement of the LVM and could for example also represent an oscillation in dorso-ventral closure or a tracheal wall vibration. We do not observe this small amplitude oscillation in our simulations.

4.6.1 Harmonics in bird song

As discussed in Section 2.3.3, the membrane and string based mathematical models were initially also implemented because they provided a mechanical basis to explain the almost pure-tone sounds generated by many birds (Casey and Gaunt, 1985; Fletcher, 1988, 1989). However, sound seems to be predominantly made by colliding syringeal labia (songbirds: Larsen and Goller, 2002) or LVMs (Goller and Larsen, 1997b) just as in humans, other mammals or geckos (Paulsen, 1967). As such, the syrinx produces a harmonically rich signal. Recent papers provide an explanation how many birds reduce

the higher harmonics and radiate their typical, almost pure tone sounds. The upper vocal tract acts as a filter in ring doves (Beckers et al., 2003a) and is actively tuned to enhance the fundamental frequency of song during vocalisation in some songbirds (Riede et al., 2006) and doves (Fletcher et al., 2004; Riede et al., 2004).

As discussed in Chapter 3 and Appendix B, our model suggests that the geometry and rest position of the syrinx can also influence the harmonic spectra drastically (Fig. 4.7). Simulations with constant rest areas (see Chapter 3) demonstrate collision-free oscillations even at high pressures (3.0 kPa = 30 $cm H_2O$), mainly because the pressure acts only on the first mass via the lower plate. This leads to an oscillation pattern characterized by weak harmonics compared with standard two-mass models, where avoidance of collisions can be obtained only at pressure values slightly above the onset pressure. During the coo simulation, the syrinx is in a persistent convergent shape ($\frac{a_1}{a_2} > 1$, Fig. 4.4a). As such, the suprasyringeal pressure has little effect on the masses, which might lead to reduced source-tract interaction. This again suggests that filter mechanisms in doves are closer to those in human phonation (Beckers et al., 2003a) in contrast to other possible explanations, such as the soprano-model.

4.6.2 Control of sound production

We can conclude that the complex interplay of bronchial pressure, syringeal muscle force generation and transmural pressure all affect precise amplitude and frequency control. In the case of ring doves, the syringeal muscles have the most profound effect on both amplitude modulation (AM) and frequency modulation (FM). Nevertheless, a separated physiological correlate with amplitude or fundamental frequency of the produced sound (Gaunt et al., 1982; Beckers et al., 2003b) seems not to be the case. Even bronchial pressure directly affects the fundamental frequency (Fig. 4.6d, 4.6e). The implementation of a muscle model to calculate force produced by the muscle instead of EMG correlates, provides explanation for the fast frequency rise and fall at the on and offset of each trill element.

Our model suggests that the control of frequency modulation during sound production is relatively independent of a pressure difference between air sacs compared to direct muscular control in ring doves. Brackenbury (1972) shows that a small caudo-cranial pressure gradient exists in the air sac system. This gradient plays an important role for the gas exchange during ventilation. He provides the only directly measured differential pressure between air sacs, and demonstrates that the difference between interclavicular air sac (ICAS) and anterior thoracic air sac is negligible, and that the difference between ICAS and abdominal air sac is only one tenth of the parent pressure. During clucking in a chicken the maximal pressure in the ICAS is 1.0 kPa (=10 $cm H_2O$), while the differential with the anterior thoracic air sac maximally only mea-

sures 0.03 kPa (Brackenbury, 1972). Furthermore, the transmural pressure is very small compared to the pressure that can be exerted on the LVM by the TL muscle (1.0 kPa vs 20 kPa).

Our in silico experiments have their in vivo experimental counterparts and thus lead to testable predictions. Switching off the TL as in Fig. 4.6, equals performing a bilateral nerve cut experiment. Switching off the transmural pressure equals an experiment where the ICAS syringeal pressure is levelled, e.g. by making a hole in the bronchus. Therefore, we expect that a coo after a bilateral nerve cut exhibits similarity to Fig. 4.6d. Levelling the bronchial and ICAS pressure might be a more difficult experiment, because the structural integrity of the syrinx is easily affected.

The small FM range of 400 Hz in ring doves is limited, but might still play an important role in perception (Beckers et al., 2001). For human listeners, an illustrative example is the two-tone vocalisation of the Common Cuckoo (*Cuculus canorus*): the two tones are distinct, but their fundamental frequency (respectively 650 and 550 Hz) differs only 100 Hz. Although many non-oscine's vocalisations may have a limited frequency range, almost every family harbours species capable of considerable FM sweeps (i.e., with ranges in the frequency modulation up to approximately 8 kHz; Bergman and Helb, 1982; Kroodsma, 2005) or intricate song (Ficken et al., 2000).

The classification of the songbirds and the mechanisms explaining the radiation of the group into the most diverse and numerically dominant avian order are still under debate (Ericson et al., 2003; Barker et al., 2004). The evolution of a complex syrinx and the ability of vocal learning is used to explain this radiation (Baptista and Trail, 1992; Raikow, 1986). The songbird syrinx design is highly conserved in evolution (Ames, 1971; King, 1989) and has 6-8 pairs of muscles (King, 1989). However, there seems to be no correlation between syringeal and song complexity (Gaunt, 1983; Baptista and Trail, 1992). Some syringeal muscles have been correlated with fundamental frequency or amplitude of sound (Goller and Suthers, 1996a,b; Suthers et al., 1999), but the biomechanical effects of each individual muscle, let alone combinations of differently recruited muscles, are far from understood. Our model suggests that independent control of song characteristics is not possible with the simple syringeal morphology of the ring dove, as suggested before by Gaunt (1983). We speculate that songbirds evolved a syrinx design that allowed for the independent control of sound parameters. Uncoupling the control of different sound parameters as an evolutionary key innovation would lead to a tremendously increase in the possible variation of song, which provides an additional explanation for the rapid diversification and speciation of songbirds.

5

Outlook

In this thesis, we have done a first step towards more realistic modelling of the avian syrinx. Our quantitative approach represents a first attempt into the delicate field of the quantitative biology; biomechanics is founded on the same basic principles as ordinary mechanics, but fewer idealizations can be made in analysis (Titze, 1994). Therefore, much of biomechanics is still at a descriptive level of analysis, although the quantification of the material properties of body tissues on the basis of molecular structure has been subject in recent years of considerable progresses. In humans, and even more in birds, one major problem of quantitative approaches in biomechanics is the accessibility of the material under investigation. Moreover, in birds only in last years the understanding of the mechanisms underlying sound production advanced very rapidly (Elemans et al., 2003).

The trapezoidal model, although simplified in many aspects, already represents a valid model to study quantitatively the contribution of different siryngeal muscles and to generate new experimentally testable hypotheses (see Chapter 4). However, the contribution of new quantitative experimental data could lead to several improvements which include:

- a more detailed description of the physiology
- the implementation of the force distribution (cf. Lous et al., 1998) rather than the simple midline assumption
- the aerodynamic description of the flow (cf. Pelorson et al., 1994)
- a more detailed quantitative investigation of the syringeal muscles. In particular (see Appendix C):
    1. the study of a weighted contribution of P_t and P_{TL}

② the study of possible further effects on labia positions and tissue elasticity (e.g., the musculus sternotrachealis ST)

Among birds, songbirds, which learn their song from a tutor, provide an excellent and widely used model system for sensorimotor learning and human speech acquisition (Doupe and Kuhl, 1999), and therefore mathematical modelling of the songbirds syrinx can help to build a bridge between the behavioral output (song) and the complex representation and execution of vocal communication at the level of the brain (Laje et al., 2002; Laje and Mindlin, 2002, 2005; Gardner et al., 2001; Mindlin et al., 2003; Mindlin and Laje, 2005). We presented in this thesis the rescaled two-mass model to study pressure onset and avoidance of collisions in the songbirds syrinx. Two-mass models (e.g., Fee et al., 1998; Fee, 2002) have been relatively seldom applied to study the songbirds syrinx, although recent insights (Goller and Larsen, 1997a,b, 2002) encourage possible two-mass model approaches in the study of the mechanisms underlying sound production in songbirds. Certainly, in account of the complex structure of the songbirds syrinx, the rescaled two-mass model presented in this thesis needs to implement a more detailed description of the physiology.

However, even with our low-order model, investigation of the respective contributions of the two labia to songbird sound production remains a challenging experimental task: although quantitative data do not exist, the lateral labium LL appears to have a larger mass than the medial labium ML (Goller and Larsen, 2002). Such mechanical asymmetries underlie nonlinear acoustic phenomena in humans (e.g., Steinecke and Herzel, 1995).

Moreover, the medial tympaniform membrane (MTM) could be incorporated in the rescaled two-mass model (e.g., via a third small mass - or membrane - attached to the mass m_1) to test the role of the MTM in songbirds, which is still debated: it could play a role in the observed wave-like motion of the composite MTM and ML structure (Fee, 2002) which has been shown to be critical for the efficient generation of sound in human larynx (Titze, 1988), it may play a distinct role especially in the generation of high frequency sounds (Goller and Larsen, 2002), or could be responsible for he suppression of higher harmonics by decreasing the amount of collisions between the masses (see Chapter 4).

In both models, we studied pressure onset, control of harmonic overtones and "registers" of the sound radiated by the birds syrinx in absence of source-tract coupling. However, in many birds vocalizations the filter of the trachea seems to be relevant (Fletcher and Tarnopolsky, 1998; Riede et al., 2004, 2006) in the production of pure tone sounds and complex vocalizations: even disregarding complex labium structure, nonlinear effects can be obtained if acoustic feedback is considered (Laje and Mindlin, 2005).

In addition to labium structure and source-tract coupling, the interaction between the two sound sources is also a possible dynamical origin of complexity (i.e., nonlinear phenomena): the ICAS (which ramifies between the two MTM's) and the songbird's pessulus (a cartilaginous tissue connecting the medial walls of the bronchi) could induce pressure connections or structural coupling between the two syringeal sources, respectively (Nowicki and Capranica, 1986a,b). A two-mass model approach with acoustic source-source and source-tract coupling could be investigated to gain insights into the mechanisms of sound production in songbirds.

Finally, an extremely challenging issue could be to investigate how some birds can imitate human speech. It has been suggested that, analogous to human speech production, tongue movements in parrot vocalisations modulate formant characteristics independently from the vocal source (Beckers et al., 2004). Moreover, Klatt and Stefanski (1974) have shown that in trained Indian Hill mynah bird, some vowel spectra have clear formant patterns but many have multiple resonances and formant splitting effects that cannot be accounted for in terms of an acoustic tube model of formant generation, as in humans (Fant, 1960; Titze, 1994).

Bibliography

P. L. Ames. The morphology of the syrinx in passerine birds. *Bull. Peabody. Mus. Nat. Hist.*, 37:1–194, 1971.

D. K. Arrowsmith and C. M. Place. *An introduction to dynamical systems*. Cambridge University Press, 1990.

M. R. Ballintijn and C. ten Cate. Sex differences in the vocalizations and syrinx of the collared-dove (*Streptopelia decaocto*). *Auk*, 144:22–39, 1997.

M. R. Ballintijn and C. ten Cate. Sound production in the collared dove: a test of the 'whistle' hypothesis. *J. Exp. Biol.*, 201:1637–1649, 1998.

M. R. Ballintijn, C. ten Cate, E. W. Nuijens, and H. Berkhoudt. The syrinx of the collared dove (*Streptopelia decaocto*): Structure, inter-individual variation and development. *Netherlands Journal of Zoology*, 45:455–479, 1995.

L. F. Baptista and P. W. Trail. The role of song in the evolution of passerine diversity. *Systematic Biology*, 41:242–247, 1992.

F. K. Barker, A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. Phylogeny and diversitification of the largest avian radiation. *Proc. Natl. Acad. Sci.*, 101:11040–11045, 2004.

G. J. L. Beckers, B. M. A. Goossens, and C. ten Cate. Perceptual salience of acoustic differences between conspecific and allospecific vocalisations in african collared-doves. *Anim. Behav.*, 62:511–518, 2001.

G. J. L. Beckers, R. A. Suthers, and C. ten Cate. Pure-tone birdsong by resonance filtering of harmonic overtones. *Proc. Natl. Acad. Sci. USA*, 100:7372–7376, 2003a.

G. J. L. Beckers, R. A. Suthers, and C. ten Cate. Mechanisms of frequency and amplitude modulation in ring dove song. *J. Exp. Biol.*, 206:1833–1843, 2003b.

G. J. L. Beckers, B. S. Nelson, and R. A. Suthers. Vocal-tract filtering by lingual articulation in a parrot. *Current Biology*, 14:1592–1597, 2004.

P. Berge, Y. Pomeau, and C. Vidal. *Order within chaos*. Wiley-Interscience, 1986.

H-H. Bergman and H-W. Helb. *Stimmen der V ogel Europas*. BLV Verlagsgesellschaft., Germany, 1982.

D. A. Berry, H. Herzel, I. R. Titze, and B. H. Story. Bifurcations in excised larynx experiments. *J. Voice*, 10:129–138, 1996.

C. D. Bertram. Flow phenomena in floppy tubes. *Contemp. Physics*, 45:45–60, 2004.

C. D. Bertram and T. J. Pedley. A mathematical model of unsteady collapsible tube behaviour. *J. Biomech.*, 15:39–50, 1982.

B. P. Bogert and G. E. Peterson. The acoustic of speech. In L. E. Travis, editor, *Handbook of speech pathology*, pages 109–173. Appleton-Century-Crofts, New York, 1957.

D. F. Boggs, P. J. Butler, and S. Wallace. Differential air-sac pressures in diving tufted ducks, anthya fuligula. *J. Exp. Biol.*, 201:2665–2668, 1998.

D. F. Boggs, R. V. Baudinette, P. B. Frappell., and P.J. Butler. The influence of locomotion on air-sac pressure in little penguins. *J. Exp. Biol.*, 204:3581–3586, 2001.

J. H. Brackenbury. Lung-air-sac anatomy and respiratory pressures in the bird. *J. Exp. Biol.*, 57:543–550, 1972.

J. H. Brackenbury. Aeroacoustics of the vocal organ of birds. *J. Theoret. Biol.*, 81: 341–349, 1979.

J. H. Brackenbury. The structural basis of voice production and its relationship to sound characteristics. In D. E. Kroodsma and E. H. Miller, editors, *Acoustic communication in birds*, volume 1, pages 53–73. Academic Press, New York, 1982.

E. F. Brittan-Powell, R. J. Dooling, and S. M. Farabaugh. Vocal development in the budgerigar (*Melopsittacus undulatus*): Contact call learning. *Journal of Comparative Psychology*, 111:226–241, 1997a.

E. F. Brittan-Powell, R. J. Dooling, O. N. Larsen, and J. T. Heaton. Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*). *J. Acoust. Soc. Amer.*, 101:578–589, 1997b.

A. M. Carroll and P. C. Wainwright. Muscle function and power during suction feeding in largemouth bass (micropterus salmoides). *Comp. Biochem. Physiol. A*, 143:389–399, 2006.

R. M. Casey and A. S. Gaunt. Theoretical models of the avian syrinx. *J. Theoret. Biol.*, 116:45–64, 1985.

D. R. Chamberlain, W. B. Gross, G. W. Cornwell, and H. S. Mosby. Syringeal anatomy in the common crow. *Auk*, 85:244–252, 1968.

A. Doupe and P. Kuhl. Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci.*, 22:567–631, 1999.

H. R. Duncker. The lung air sac system of birds. a contribution to the functional anatomy of the respiratory apparatus. *Ergeb. Anat. Entwicklungsgesch.*, 45:7–171, 1971.

J. P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.*, 57(3):617–656, Jul 1985.

C. P. H. Elemans. *How do birds sing? sound analysis - mechanical modelling - muscular control*. PhD thesis, Experimental Zoology Group, Wageningen University, Wageningen, The Netherlands, 2004.

C. P. H. Elemans, O. N. Larsen, M. R. Hoffmann, and J. L. van Leeuwen. Quantitative modeling of the biomechanics of the avian syrinx. *Animal Biol.*, 53:183–193, 2003.

C. P. H. Elemans, L. Y. Spierts, U. K. Mueller, J. L. van Leeuwen, and F. Goller. Superfast muscles control doves trill. *Nature*, 431:146–146, 2004.

C. P. H. Elemans, L. Y. Spierts, M. Hendriks, H. Schipper, U. K. Mueller, and J. L. van Leeuwen. Syringeal muscles fit the trill in ring doves (*Streptopelia risoria*)s. *J. Exp. Biol.*, 209:965–977, 2006.

C. P. H. Elemans, R. Zaccarelli, and H. Herzel. Biomechanics and control of vocalization in a non-songbird. *J. Royal Soc. Interface*, November 2007. doi: 10.1098.

P. G. P. Ericson, M. Irestedt, and U. S. Johansson. Evolution, biogeography, and patterns of diversification in passerine birds. *J. Avian Biol.*, 34:3–15, 2003.

B. Ermentrout. *XPPAUT5.0 - the differential equations tool*, January 2001. URL `http://www.math.pitt.edu/~bard/bardware/xpp_doc.pdf`.

G. Fant. *Acoustic theory of speech production*. The Hague: Mouton & Company, 1960.

M. S. Fee. Measurement of the linear and nonlinear mechanical properties of the oscine syrinx: Implications for function. *J. Comp. Physiol. A.*, 188:829–839, 2002.

M. S. Fee, B. Shraiman, B. Pesaran, and P. P Mitra. The role of nonlinear dynamics of the syrinx in the vocalization of a songbird. *Nature*, 395:67–71, 1998.

J. E. Ffowcs Williams and D. J. Lovely. Sound radiation into uniformly flowing fluid by compact surface vibration. *J. Fluid. Mech.*, 71:689–700, 1975.

M. S. Ficken, K. M. Rusch, S. J. Taylor, and D. R. Powers. Blue-throated hummingbird song: a pinnacle of nonoscine vocalization. *Auk*, 117:120–128, 2000.

W. T. Fitch, J. Neubauer, and H. Herzel. Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, 63: 407–418, 2002.

J. L. Flanagan and L. Landgraf. Self-oscillating source for vocal-tract synthesizers. *IEEE Transactions on Audio and Electroacoustics*, AU-16:57–64, 1968.

N. H. Fletcher. Mode locking in nonlinearly excited inharmonic musical oscillators. *J. Acoust. Soc. Amer.*, 64:1566–1569, 1978.

N. H. Fletcher. Bird song - a quantitative acoustic model. *J. Theoret. Biol.*, 135:455–481, 1988.

N. H. Fletcher. Acoustics of bird song - some unresolved problems. *Comm. Theor. Biol.*, 1:237–251, 1989.

N. H. Fletcher. *Acoustic Systems in Biology*. Oxford University Press, Oxford, New York, 1992.

N. H. Fletcher and T. D. Rossing. *The Physics of Musical Instruments*. Springer, 2005.

N. H. Fletcher and A. Tarnopolsky. Acoustics of the avian vocal tract. *J. Acoust. Soc. Amer.*, 105(1):35–49, 1998.

N. H. Fletcher, T. Riede, G. J. L. Beckers, and R. A. Suthers. Vocal tract filtering and the “coo” of doves. *J. Acoust. Soc. Amer.*, 116:3750–3756, 2004.

T. Frank, I. Walter, A. Probst, and H.E. K onig. Histological aspects of the syrinx of the male mallard (anas plathyrhynchos). *Anat. Histol. Embryol.*, 35:396–401, 2006.

H. Froehling, J. P. Crutcheld, D. Farmer, N. H. Packard, and R. Shaw. On determining the dimension of chaotic flows. *Physica 3D*, 3(3):605–617, 1981.

T. J. Gardner, G. Cecchi, M. Magnasco, R. Laje, and G. B. Mindlin. Simple motor gestures for birdsongs. *Phys. Rev. Lett.*, 87(20):208101, Oct 2001.

A. S. Gaunt. An hypothesis concerning the relationship of syringeal structure to vocal abilities. *Auk*, 100:853–862, 1983.

A. S. Gaunt and S. L. L. Gaunt. Syringeal structure and avian phonation. In R. F. Johnston, editor, *Current Ornithology*, volume 2, pages 213–245. Plenum Press, New York, 1985.

A. S. Gaunt, S. L. L. Gaunt, and R. M. Casey. Syringeal mechanism reassessed: evidence from *Streptopelia*. *Auk*, 99:474–494, 1982.

F. Goller and O. N. Larsen. New perspectives on mechanism of sound generation in songbirds. *J. Comp. Physiol.*, 188:841–850, 2002.

F. Goller and O. N. Larsen. A new mechanism of sound generation in songbirds. *Proc. Natl. Acad. Sci. USA*, 94:14787–14791, 1997a.

F. Goller and O. N. Larsen. *In situ* biomechanics of the syrinx and sound generation in pigeons. *J. Exp. Biol.*, 200:2165–2176, 1997b.

F. Goller and R. A. Suthers. Role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J. Neurophysiol.*, 75:867–876, 1996a.

F. Goller and R. A. Suthers. Role of syringeal muscles in controlling the phonology of bird song. *J. Neurophysiol.*, 76:287–300, 1996b.

D. Goodwin. Pigeons and doves of the world. London: British Museum of Natural History, 1983.

C. H. Greenewalt. *Bird Song: Acoustics and Physiology*. Smithsonian Institution Press, Washington, D. C., 1968.

J. B. Grotberg and O. E. Jensen. Biofluid mechanics in flexible tubes. *Ann. Rev. Fluid Mech.*, 36:121–147, 2004.

J. Guckenheimer. Numerical analysis of dynamical system, October 1999. URL `www.math.cornell.edu/~\gucken/PDF/algorithms.pdf`.

J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, 1983.

H. Hatzikirou, W. T. Fitch, and H. Herzel. Voice instabilities due to source-tract interactions. *Acta Acustica united with Acustica*, 92:468–475, 2006.

H. Herzel. Bifurcations and chaos in voice signals. *Appl. Mech. Rev.*, 46:399–413, 1993.

H. Herzel and C. Knudsen. Bifurcations in a vocal fold model. *Nonlinear Dynamics*, 7 (1):53–64, January 1995.

H. Herzel and J. Wendler. Evidence of chaos in phonatory samples. In *EUROSPEECH*, pages 263–266, Genova, 1991.

H. Herzel, D. A. Berry, I. R. Titze, and M. Saleh. Analysis of vocal disorders with methods from nonlinear dynamics. *Journal of Speech and Hearing Research*, 37: 1008–1019, 1994.

H. Iijima, N. Mild, and N. Nagai. Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis. *IEEE Trans. Signal Proc.*, 40:2125–2135, 1992.

K. Ishizaka and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell. Syst. Tech. J.*, 51:1233–1268, 1972.

K. Ishizaka and M. Matsudaira. Fluid mechanical considerations of vocal cord vibration. *SCRL Monograph 8*, 1972.

N. Isshiki, M. Tanabe, and M. Sawada. Arytenoid adduction for unilateral vocal cord paralysis. *Arch. Otolaryngol.*, 104:555–558, 1978.

E. Atlee Jackson. *Perspectives of nonlinear dynamics*. Cambridge University Press, Melbourne, 1989.

J. J. Jiang, Y. Zhang, and J. Stern. Modeling of chaotic vibrations in symmetric vocal folds. *J. Acoust. Soc. Amer.*, 110:2120–2128, 2001.

R. D. Kent. Imitation of synthesized vowels by preschool children. *J. Acoust. Soc. Amer.*, 63:1193–1198, 1978.

R. D. Kent. Isovowel lines for the evaluation of vowel formant structure in speech disorders. *Journal of Speech and Hearing Disorders*, 44:513–521, 1979.

A. S. King. Functional anatomy of the syrinx. In A. S. King and J. McLelland, editors, *Form and Function in Birds*, pages 105–192. Academic Press, New York, 1989.

D. H. Klatt and R. A. Stefanski. How does a mynah bird imitate human speech? *J. Acoust. Soc. Amer.*, 55:822–832, 1974.

M. Kob. *Physical Modeling of the Singing Voice*. Logos Verlag Berlin, Gubener strasse 47, D-10243 Berlin (Germany), 2002.

P. V. Komi. Stretch-shortening: a powerful model to study normal and fatigued muscle. *J. Biomech.*, 33:1197–1206, 2000.

D. E. Kroodsma. *The singing life of birds: The art and science of listening to birdsong*. Houghton Mifflin Company, Boston, USA, 2005.

R. Laje and G. B. Mindlin. Diversity within a birdsong. *Physical Review Letters*, 89(28): 288102(4), 2002.

R. Laje and G. B. Mindlin. Modeling source-source and source-filter acoustic interaction in birdsong. *Physical Review E*, 72:1–11, 2005.

R. Laje, T. J. Gardner, and G. B. Mindlin. Continuous model for vocal fold oscillations to study the effect of feedback. *Physical Review E*, 64:056201, 2001.

R. Laje, T. J. Gardner, and G. B. Mindlin. Neuromuscular control of vocalizations in birdsong: A model. *Physical Review Letters E*, 65:051921(8), 2002.

O. N. Larsen and F. Goller. Direct observation of syringeal muscle function in songbirds and a parrot. *J. Exp. Biol.*, 205:25–35, 2002.

O. N. Larsen and F. Goller. Role of syringeal vibrations in bird vocalisations. *Proc. Roy. Soc. Lond.B.*, 266:1609–1615, 1999.

J. Liljenkrants. Numerical simulations of glottal flow. In *EUROSPEECH*, pages 255–258, Genova, 1991.

N. J. C. Lous, G. C. Hofmans, R. N. J. Veldhuis, and A. Hirschberg. A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design. *Acta Acustica*, 84:1135–1150, 1998.

J. C. Lucero. Dynamics of the two-mass model of the vocal folds: Equilibria, bifurcations, and oscillation region. *J. Acoust. Soc. Amer.*, 94:3104–3111, 1993.

J. C. Lucero. A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset. *J. Acoust. Soc. Amer.*, 105(1):423–431, 1999.

W. Mende, H. Herzel, and K. Wermke. Bifurcations and chaos in newborn infant cries. *Physics Letters A*, 145:418–424, apr 1990.

P. Mergell and H. Herzel. Modelling biphonation - the role of the vocal tract. *Speech Communication*, 22:141–154, 1997.

P. Mergell, H. Herzel, T. Wittenberg, M. Tigges, and U. Eysholdt. Phonation onset: vocal fold modeling and high-speed glottography. *J. Acoust. Soc. Amer.*, 104:464–470, 1998.

P. Mergell, H. Herzel, and W. T. Fitch. Modeling the role of nonhuman vocal membranes in phonation. *J. Acoust. Soc. Amer.*, 105:2020–2028, 1999.

P. Mergell, H. Herzel, and I. R. Titze. Irregular vocal fold vibration - high-speed observation and modeling. *J. Acoust. Soc. Amer.*, 108:2996–3002, 2000.

G. B. Mindlin and R. Laje. *The Physics of Birdsong*. Springer, Heidelberg, Germany, 2005.

G. B. Mindlin, T. J. Gardner, F. Goller, and R. A. Suthers. Experimental support for a model of birdsong production. *Phys. Rev. E*, 68(4):041908, 2003.

J. Neubauer. *Nonlinear Dynamics of the Voice: Bifurcations and Mode Analysis of Complex Spatio-Temporal Signals*. PhD thesis, Humboldt University, Berlin (Germany), 2004.

J. Neubauer, M. Edgerton, and H. Herzel. Nonlinear phenomena in contemporary vocal music. *J. Voice*, 18(1):1–12, 2004.

R. W. Norman and P. V. Komi. Electromechanical delay in skeletal muscle under normal movement conditions. *Acta Physiol. Scand.*, 106:241–248, 1979.

F. Nottebohm. The origins of vocal learning. *Am. Nat.*, 106:116–140, 1972.

F. Nottebohm. Phonation in orange-winged amazon parrot, *Amazona amazonica*. *J. Comp. Physiol. A.*, 108:157–170, 1976.

S. Nowicki. Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature*, 235:53–55, 1987.

S. Nowicki and R. R. Capranica. Bilateral syringeal interaction in the production of an oscine bird sound. *Science*, 231:1297–1299, 1986a.

S. Nowicki and R. R. Capranica. Bilateral syringeal coupling during phonation of a songbird. *The Journal of Neuroscience*, 6(12):3595–3610, 1986b.

S. Nowicki, P. Marler, A. Maynard, and S. Peters. Is the tonal quality of birdsongs learned? *Ethology*, 90:225–235, 1992.

K. Paulsen. *Das Prinzip der Stimmbildung in der Wirbeltierreihe und beim Menschen*. Akademische Verlagsgesellschaft, Frankfurt am Main, Germany, 1967.

X. Pelorson, A. Hirschberg, R. R. Van Hassel, A. P. J. Wijnands, and Y. Auregan. Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. application to a modified two-mass model. *J. Acoust. Soc. Amer.*, 96: 3416–3431, 1994.

G. E. Peterson and H. L. Barney. Control methods used in a study of vowels. *J. Acoust. Soc. Amer.*, 24:175–184, 1952.

R. J. Raikow. Why are there so many kinds of passerine birds? *Syst. Zool.*, 35:255–259, 1986.

T. Riede, G. J. L. Beckers, W. E. Blevins, and R. A. Suthers. Inflation of the esophagus and vocal tract filtering in ring doves. *J. Exp. Biol.*, 207:4025–4036, 2004.

T. Riede, R. A. Suthers, N. H. Fletcher, and W. E. Blevins. Songbirds tune their vocal tract to the fundamental frequency of their song. *Proc. Natl. Acad. Sci.*, 103:5543–5548, 2006.

M. Rothenberg. Acoustic interaction between the glottal source and the vocal tract. In K. N. Stevens and M. Hirano, editors, *Vocal Fold Physiology*, pages 305–323. University of Tokyo Press, 1981.

R. Scherer and I. R. Titze. Pressure-flow relationships in a model of the laryngeal airway with a diverging glottis. In D. M. Bless and J. H. Abbs, editors, *Vocal fold physiology: Contemporary research and clinical issues*, pages 179–193. College-Hill Press, San Diego, 1983.

D. Sciamarella and C. D'Alessandro. On the acoustic sensitivity of a symmetrical two-mass model of the vocal folds to the variation of control parameters. *Acta Acustica united with Acustica*, 90:746–761, 2004.

D. Sciamarella and C. D'Alessandro. Stylization of glottal-flow spectra produced by a mechanical vocal-fold model. *Proc. of the 9th European Conference on Speech Communication and Technology, Interspeech*, pages 2149–2152, 2005.

M. E. Smith, G. S. Berke, B. R. Gerrat, and J. Kreiman. Laryngeal paralyses: Theoretical considerations and effects on laryngeal vibration. *J. Speech. Hear. Res.*, 35: 545–554, 1992.

I. Steinecke and H. Herzel. Bifurcations in an asymmetric vocal-fold model. *J. Acoust. Soc. Amer.*, 97:1874–1884, 1995.

B. H. Story. Chaos in voice and speech. Research Report: 03:590, 1991.

B. H. Story and I. R. Titze. Voice simulation with a body-cover model of the vocal folds. *J. Acoust. Soc. Amer.*, 97:1249–1260, 1995.

S. H. Strogatz. *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Perseus Books, Cambridge MA, 1994.

R. A. Suthers. Contributions to birdsong from the left and right sides of the intact syrinx. *Nature*, 347:473–477, 1990.

R. A. Suthers and S. A. Zollinger. Producing song: The vocal apparatus. *New York Acad. Sci.*, 1016:109–129, 2004.

R. A. Suthers, F. Goller, and R. S. Hartley. Motor dynamics of song production by mimic thrushes. *J. Neurobiol.*, 25:917–936, 1994.

R. A. Suthers, F. Goller, and C. Pytte. The neuromuscular control of birdsong. *Philos Trans R Soc Lond B*, 354:927–939, 1999.

F. Takens. Detecting strange attractors in turbulence. *Physical Review Letters*, 79(8): 1475–1478, 1981.

I. R. Titze. *Principles of Voice Production*. Prentice-Hall, Englewood Cliffs, 2002.

I. R. Titze. Source-filter interaction in speaking and singing is nonlinear. *Echoes - The newsletter of The Acoustical Society of America*, 17:1–3, 2007.

I. R. Titze. On the mechanics of vocal-fold vibration. *J. Acoust. Soc. Amer.*, 60(6): 1366–1380, 1976.

I. R. Titze. The physics of small-amplitude oscillation of the vocal folds. *J. Acoust. Soc. Amer.*, 83:1536–1552, 1988.

I. R. Titze. *Principles of Voice Production*. Prentice-Hall, Englewood Cliffs, 1994.

I. R. Titze and R. Scherer. Mechanisms of sustained oscillation of the vocal folds. In I. Titze and R. Scherer, editors, *Vocal fold physiology: Biomechanics, acoustics, and phonatory control*, pages 349–357. Denver Center for the Performing Arts, Denver, 1983.

I. R. Titze, R. Baken, and H. Herzel. Evidence of chaos in vocal fold vibration. In I. R. Titze, editor, *Vocal Fold Physiology: New Frontier in Basic Science*, pages 143–188. Singular Publishing Group, San Diego, 1993.

I. Tokuda, T. Riede, J. Neubauer, M. J. Owren, and H. Herzel. Nonlinear analysis of irregular animal vocalizations. *J. Acoust. Soc. Amer.*, 111(6):2908–2919, 2002.

I. Tokuda, J. Horáček, J. G. Švec, and H. Herzel. Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments. *J. Acoust. Soc. Amer.*, 122(1):519–531, 2007.

J. W. van den Berg. Myoelastic aerodynamic theory of voice production. *Journal of Speech and Hearing Research*, 1:227–244, 1958.

J. W. van den Berg, J. T. Zantema, and P. Jr. Doornenbal. On the air resistance and the bernoulli effect of the human larynx. *J. Acoust. Soc. Amer.*, 29:626–631, 1957.

R. W. Warner. The syrinx in the family columbidae. *J. Zool. Lond.*, 166:285–390, 1972.

J. M. Wild. Functional anatomy and neural pathways contributing to the control of song production in birds. *Eur. J. Morph.*, 35:303–325, 1997.

I. Wilden, H. Herzel, G. Peters, and G. Tembrock. Subharmonics, biphonation, and deterministic chaos in mammal vocalisation. *Bioacoustics*, 9:171–196, 1998.

H. Williams, J. Cynx, and F. Nottebohm. Timbre control in zebra finch (*Taeniopygia guttata*) song syllables. *J. Comp. Psychol.*, 103:366–380, 1989.

A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano. Determining lyapunov exponent from a timeseries. *Phys D*, 16:285–317, 1985.

R. Zaccarelli, C. P. H. Elemans, W. T. Fitch, and H. Herzel. Two-mass models of the bird syrinx. In *Proceed. 4th MAVEBA Workshop*, volume 92, pages 47–50, Firenze, 2005. Univeristy Press.

R. Zaccarelli, C. P. H. Elemans, W. T. Fitch, and H. Herzel. Modelling bird songs: Voice onset, overtones and registers. *Acta Acustica united with Acustica*, 92:741–748(8), September/October 2006. URL `http://www.ingentaconnect.com/content/dav/aaua/2006/00000092/00000005/art00010`.

Y. Zhang and J. J. Jiang. Chaotic vibrations of a vocal fold model with a unilateral ploy p. *J. Acoust. Soc. Amer.*, 115(4):1266–1269, 2003.

A

Trapezoidal model: geometrical definitions and calculation of forces

A.1 Derivation of geometrical quantities

We present here a detailed description of the geometry of the trapezoidal model and the calculation of forces acting on the masses. Our results are implement in MATLAB code to provide easy visualization of the dynamic changes of the model.

In the syringeal frontal plane, we denote the horizontal axis as x and the vertical as z (Fig. A.1). As in previous models (Ishizaka and Flanagan, 1972; Steinecke and Herzel, 1995), $x_1 = x_1(t)$, $x_2 = x_2(t)$ in eq. 3.1-3.4 on p. 47 will indicate the displacement of mass m_1, m_2, while x_{01}, x_{02} will denote the rest position of the two masses (we suppose no vertical motion along the z axis).

Points On the plane xz, we define the point T_i as the position of m_i, $i \in \{1,2\}$, T_0, T_3 as the boundary points of the LTM on the tracheal walls, and $T_M \in \gamma_2$ as the point located

at the γ_2 height midline. The geometrical coordinates of these points are:

$$
\begin{aligned}
T_0 &= (w, 0) \\
T_1 &= (x_{01} + x_1, d_1) \\
T_M &= \left(\frac{x_{01} + x_1 + x_{02} + x_2}{2}, d_1 + \frac{d_2}{2} \right) \\
T_2 &= (x_{02} + x_2, d_1 + d_2) \\
T_3 &= (w, d_1 + d_2 + d_3)
\end{aligned}
$$

We denote as z_0, z_1, z_M, z_2, z_3 the ordinates of T_0, T_1, T_M, T_2, T_3, respectively.

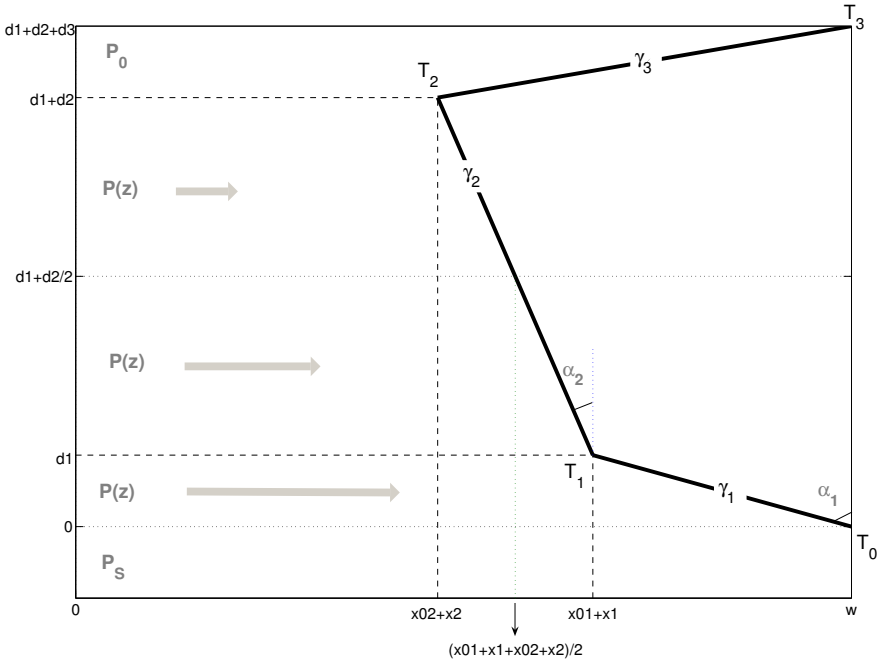


Figure A.1: 2-D view of the trapezoidal model in the frontal plane xz. Note the comparative lengths of the arrows labeled $P(z)$ to indicate the action of the Bernoulli pressure $P(z)$ at height z (eq. A.1).

Lines The line segment γ_i, $i \in \{1,2,3\}$ will identify the i$^{\text{th}}$ mass-less plate of every LTM: it is defined as the part of the line passing for T_{i-1} and T_i and bounded between these two points. From the geometrical coordinates of each line[1], choosing z as the independent variable, we can get the abscissa $x = f(z)$ of a generic point T placed on

[1] Given two points $T_1 = (x_1, y_1)$ and $T_2 = (x_2, y_2)$ the equation of the (unique) line through T_1, T_2 is:

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}$$

the LTM:

$$x(z) = \begin{cases} \dfrac{x_{01}+x_1-w}{d_1} z + w & \text{if } 0 \le z \le d_1 \\ \dfrac{x_{02}+x_2-x_{01}-x_1}{d_2}(z-d_1)+(x_{01}+x_1) & \text{if } d_1 < z \le d_1+d_2 \\ \dfrac{w-x_{02}-x_2}{d_3}(z-(d_1+d_2))+(x_{02}+x_2) & \text{if } d_1+d_2 < z \le d_3 \\ 0 & \text{otherwise} \end{cases}$$

Syringeal areas In the plane xy, a_1 and a_2 will denote, as in previous models, the part of the tracheal area formed by the position of the masses m_1 and m_2. The syringeal area a_i ($i \in \{1,2\}$) reads $a_i = 2\ell(x_{0i}+x_i)$, where ℓ is the length of the syringeal lumen along the y axis (see Fig. A.2).

We define $a_{min} = \min\{a_1, a_2\}$. In the same way are defined the areas a_0, a_M, a_3 corresponding to the points T_0, T_M, T_3. Generalizing, we define syringeal area $a(z)$ at height z as follows:

$$a(z) = \begin{cases} 2\ell x(z) & \text{if } 0 \le z \le d_3 \\ 0 & \text{otherwise} \end{cases}$$

where $x(z)$ has been obtained in the former paragraph by means of LTM plate's equations. Note that, in our notation, $a_1 = a(d_1)$, $a_2 = a(d_1+d_2)$.

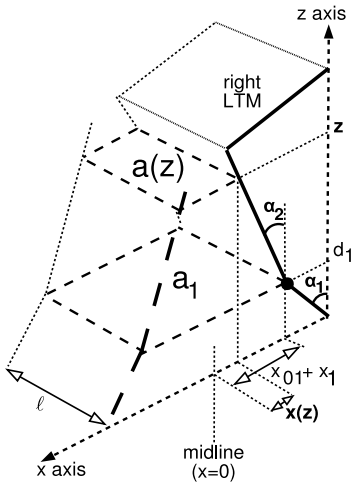


Figure A.2: Trapezoidal model in 3D view.

Collision points Collision of the LTM is geometrically equivalent to the condition $a_{min} < 0$. During collision, the property $a(z) = 0$ holds for two and only two points on the LTM plates. We define mathematically the collision point ordinate (c.p.o.) of the i$^{\text{th}}$ plate as:

$$\zeta_i = \{z_T, T = (x_T, z_T) \in \gamma_i : a(z_T) = 0\}, \quad i \in \{1,2,3\}$$

Therefore, according to the masses displacement a pair of collision point ordinates (ζ_1, ζ_2), (ζ_2, ζ_3) or (ζ_1, ζ_3) is univocally determined during collision. For each pair, we denote the minimum c.p.o. as $\zeta_{min} = \min\{\zeta_i, \zeta_j\}$, $i \neq j$.

Angles We define the angles α_1 and α_2 (see Fig. A.2), both lying on the plane xz, as follows:

- α_1: plane angle formed by the the line γ_1 and the vertical line (parallel to z axis) passing for T_0
- α_2: plane angle formed by the line γ_2 and the vertical line (parallel to z axis) passing to T_1

In the following, given a line γ_i, $i \in \{1,2\}$ we will refer as α_i as "the angle formed by γ_i".

According to the previous definitions, in Table A.1 are summarized the main geometrical quantities which will be recalled throughout this section.

points (c.p.o. =collision point ordinate)

$z_0 = 0$	T_0 ordinate
$z_1 = d_1$	T_1 ordinate
$z_M = d_1 + \frac{d_2}{2}$	T_M ordinate
$z_2 = d_1 + d_2$	T_2 ordinate
$z_3 = d_1 + d_2 + d_3$	T_3 ordinate
$\zeta_1 = \dfrac{a_0 d_1}{a_0 - a_1}$	γ_1 c.p.o.
$\zeta_2 = d_1 - \dfrac{a_1 d_2}{a_2 - a_1}$	γ_2 c.p.o.
$\zeta_3 = d_1 + d_2 - \dfrac{a_2 d_3}{a_0 - a_2}$	γ_3 c.p.o.

syringeal areas (s.a.)

$a_0 = 2\ell w$	s.a. at height z_0
$a_1 = 2\ell(x_{01} + x_1)$	s.a. at height z_1
$a_M = 2\ell(\frac{x_{01}+x_1+x_{02}+x_2}{2})$	s.a. at height z_M
$a_2 = 2\ell(x_{02} + x_2)$	s.a. at height z_2
$a_3 = 2\ell w$	s.a. at height z_3

$$a(z) = \begin{cases} \left(\dfrac{a_1 - a_0}{d_1}\right) z + a_0 & \text{if } 0 \le z \le d_1 \\ \left(\dfrac{a_2 - a_1}{d_2}\right)(z - d_1) + a_1 & \text{if } d_1 < z \le d_1 + d_2 \\ \left(\dfrac{a_0 - a_2}{d_3}\right)(z - d_1 - d_2) + a_2 & \text{if } d_1 + d_2 < z \le d_3 \end{cases} \quad \text{s.a. at height } z$$

Table A.1: Geometrical quantities.

A.2 Calculation of forces

We assume that the air stream separates from the vocal fold surface at the smallest syringeal opening area (a_{min}) to form a jet. This jet keeps the pressure above the jet separation point close to zero. According to the Bernoulli equation (see Section 2.1.2), when the syrinx is open ($a_{min} > 0$) we can define at any point in the syrinx lumen between a_0 and a_{min} the pressure at height z as

$$P(z) = P_s - \frac{\rho}{2}\frac{U^2}{a^2(z)} \tag{A.1}$$

where ρ is the air density, U is the flow (and consequently $\frac{U}{a(z)}$ is the particle velocity), $a(z)$ is the syringeal area at height z (see Table A.1) and P_s is the subsyringeal pressure.

A.2.1 Flow

Along the lines of Steinecke and Herzel (1995), we denote the suprasyringeal pressure P_{sup} as P_0 (see Fig. A.1). Setting $P_0 = 0$, the flow can be calculated from eq. A.1:

$$P_s = P_0 + \frac{\rho}{2}\frac{U^2}{a_{min}^2} \Rightarrow (P_0 = 0) \Rightarrow P_s = \frac{\rho}{2}\frac{U^2}{a_{min}^2}.$$

Therefore:

$$U = a_{min}\sqrt{\frac{2P_s}{\rho}} \tag{A.2}$$

A.2.2 Pressure at height z

Replacing U in eq. A.1 with eq. A.2 we obtain:

$$P(z) = P_s - \frac{\rho}{2a^2(z)}\left[a_{min}^2\left(\frac{2P_s}{\rho}\right)\right] = P_s - \frac{P_s a_{min}^2}{a^2(z)} = P_s\left[1 - \left(\frac{a_{min}}{a(z)}\right)^2\right] \tag{A.3}$$

Which is true only if eq. A.2 holds, i.e. if z is in the syrinx lumen between a_0 and a_{min} and $a_{min} > 0$. A general formula of $P(z)$ during open configuration can be expressed by setting $P(z) = 0$ for any point above the point where the jet separation occurs:

$$P(z) = P_s\left[1 - \left(\frac{a_{min}}{a(z)}\right)^2\right]\Theta(z_{amin} - z) \tag{A.4}$$

where $\Theta(x)$ is the *Heaviside* function:

$$\Theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

and z_{amin} is the ordinate of the point at which a_{min} is found, i.e., $z_{amin} = d_1$ if $a_{min} = a_1$, otherwise ($a_{min} = a_2$) $z_{amin} = d_1 + d_2$.

Eq. A.4 holds for an open syrinx. In case of collision ($a_{min} \leq 0$), all the points with ordinate in the interval $[0, \zeta_{\min}]$ are exposed to the subsyringeal pressure P_s, where ζ_{min} is the minimum c.p.o.. Therefore, if $a_{min} \leq 0$:

$$P(z) = P_s \Theta(\zeta_{\min} - z) \tag{A.5}$$

Combining eq. A.4 and eq. A.5, the pressure $P(z)$ at any point z in the syringeal lumen reads:

$$P(z) = \begin{cases} P_s \left[1 - \left(\dfrac{a_{min}}{a(z)} \right)^2 \right] \Theta(z_{amin} - z) & \text{if } a_{min} > 0 \\ \\ P_s \Theta(\zeta_{min} - z) & \text{if } a_{min} \leq 0 \end{cases}$$

In Table A.2 we summarize the values of $P(z)$ according to the syringeal geometry.

A.2.3 Pressure force on a plate area

By definition, the pressure force exerted on an area A and normal to A is:

$$F = \int_A P \, \mathrm{d}A$$

In our case, P is the pressure $P(z)$ calculated in the previous section, while A is the surface of the LTM on which this pressure acts (see Fig. A.3). For a generic point T at height z, the LTM surface at height z can be easily expressed as $A(z) = \ell \frac{z}{\cos\alpha}$, where α is the angle formed by the plate where T lies (Fig. A.3).

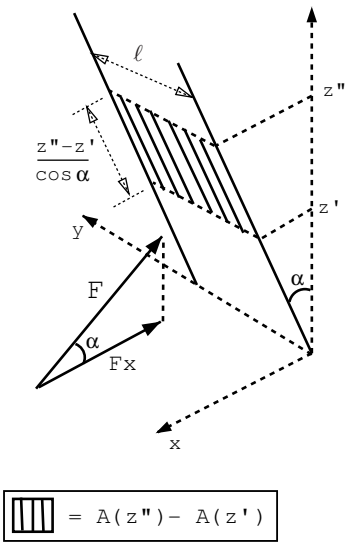


Figure A.3: F components acting on the LTM.

Therefore, if we consider two generic ordinates z' and z'' both located on the same plate γ and $z' < z''$ (see Fig. A.3), the surface area between z' and z'' is given by $A(z'') - A(z')$. Consequently, the pressure force $F(z', z'')$ acting on a plate γ between z' and z'' reads:

$$F(z', z'') = \int_{\ell \frac{z'}{\cos\alpha}}^{\ell \frac{z''}{\cos\alpha}} P(z) \, \mathrm{d}A(z) \overset{(1)}{=} \int_{z'}^{z''} P(z) \frac{\ell}{\cos\alpha} \, \mathrm{d}z = \frac{\ell}{\cos\alpha} \int_{z'}^{z''} P(z) \, \mathrm{d}z$$

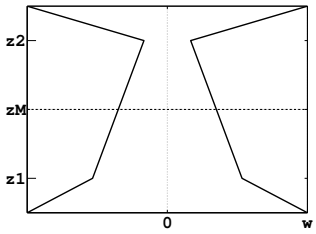
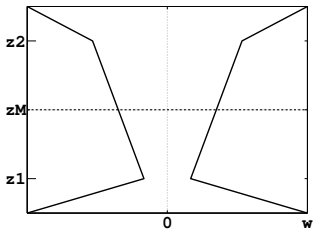
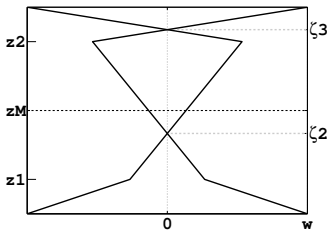
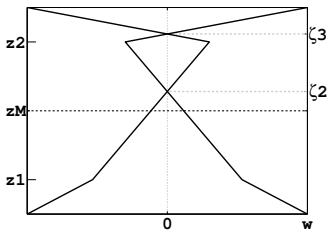
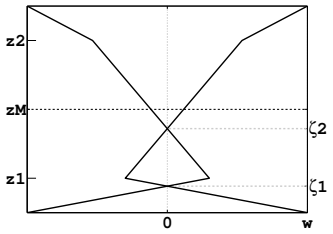
configuration	**pressure** $P(z)$
[1] convergent open 	$P(z) = \begin{cases} P_s \left[1 - \frac{a_2^2}{a^2(z)}\right] & \text{if } 0 < z < z_2 \\ 0 & \text{otherwise} \end{cases}$
[2] divergent open 	$P(z) = \begin{cases} P_s \left[1 - \frac{a_1^2}{a^2(z)}\right] & \text{if } 0 < z < z_1 \\ 0 & \text{otherwise} \end{cases}$
[3a] closed $a_1 > 0,\ a_1 \leq \lvert a_2 \rvert$  ***[3b] closed*** $a_1 > 0,\ a_1 > \lvert a_2 \rvert$ 	$P(z) = \begin{cases} P_s & \text{if } 0 < z < \zeta_2 \\ 0 & \text{otherwise} \end{cases}$ $P(z) = \begin{cases} P_s & \text{if } 0 < z < \zeta_2 \\ 0 & \text{otherwise} \end{cases}$
[4] closed $a_1 \leq 0$ 	$P(z) = \begin{cases} P_s & \text{if } 0 < z < \zeta_1 \\ 0 & \text{otherwise} \end{cases}$

Table A.2: Trapezoidal model: syrinx configurations and relative pressure values. The redundant cases [3a] and [3b] are displayed separately because they will lead to different pressure force values derived from $P(z)$ (see Table A.3)

where in (1) we used the relation $\mathrm{d}A = \frac{\ell}{\cos\alpha}\,\mathrm{d}z$.

Because no vertical motion is supposed, the real force acting on the plate area between z' and z'' which affects the motion of the masses will be the component of $F(z',z'')$ along the x direction:

$$F_X(z',z'') = F(z',z'')\cos\alpha = \ell\int_{z'}^{z''} P(z)\,\mathrm{d}z. \tag{A.6}$$

A.2.4 Pressure forces acting on the masses

For $i \in \{1,2\}$ we denote as F_i the force generated on the i$^{\text{th}}$ mass by the pressure on the syringeal lumen. We assume for simplicity (*midline assumption*):

① The pressure acting on a point z on the first plate will affect the motion of the 1$^{\text{st}}$ mass, as well as the pressure on a point z located on the *lower* half of the second plate γ_2 ($d_1 < z \leq d_1 + \frac{d_2}{2}$):

$$F_1 = F_X(z_0,z_1) + F_X(z_1,z_M) = \ell\int_0^{d_1} P(z)\,\mathrm{d}z + \ell\int_{d_1}^{d_1+\frac{d_2}{2}} P(z)\,\mathrm{d}z \tag{A.7}$$

② The pressure acting on a point z located on the upper half of the second plate γ_2 ($z > d_1 + \frac{d_2}{2}$) will affect the motion of the second mass m_2:

$$F_2 = F_X(z_M,z_2) = \ell\int_{d_1+\frac{d_2}{2}}^{d_1+d_2} P(z)\,\mathrm{d}z \tag{A.8}$$

According to the syrinx geometry, which affects the pressure $P(z)$, eq. A.7-A.8 reduce to the formulas listed in Table A.3 (cf. Table A.2); in a closed configuration, as $P(z)$ is nonzero up to the minimum c.p.o. ζ_{min} (see eq. A.5), according to the midline assumption we have $F_2 = F_X(z_M,z_2) = 0$, $F_X(z_1,z_M) = 0$ when $\zeta_{min} = \zeta_1 < z_1$ (case [4] in Table A.3) and two distinct cases when $\zeta_{min} = \zeta_2$: $\zeta_2 > \zeta_M$ (case [3b]) and $\zeta_2 \leq \zeta_M$ (case [3a]: here again $F_2 = 0$). Nevertheless, when nonzero F_1 and F_2 can be easily calculated by definition by means of eq. A.6:

$$F_X(z',z'') = \ell\int_{z'}^{z''} P(z)\,\mathrm{d}z = \ell\int_{z'}^{z''} P_s\Theta(\zeta_{min}-z)\,\mathrm{d}z = \ell P_s(\min\{\zeta_{min},z''\} - z')$$

In an open configuration, as $P(z)$ is nonzero up to z_{amin} (jet separation, see eq. A.4), we have $F_2 = F_X(z_M,z_2) = 0$, $F_X(z_1,z_M) = 0$ when $z_{amin} = z_1$ (case [2], divergent configuration). Otherwise, the components F_X whereby we calculated F_1 and F_2 read (note that we can suppose both $z',z'' < z_{amin}$, i.e. $\Theta(z_{amin}-z) = 1$ in eq. A.4):

$$F_X(z',z'') = \ell\int_{z'}^{z''} P_s\left(1 - \frac{a_{min}^2}{a(z)^2}\right)\mathrm{d}z = \ell P_s(z''-z') - \ell P_s a_{min}^2\int_{z'}^{z''}\frac{1}{a(z)^2}\,\mathrm{d}z$$

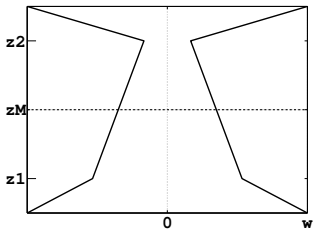
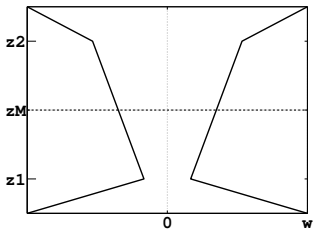
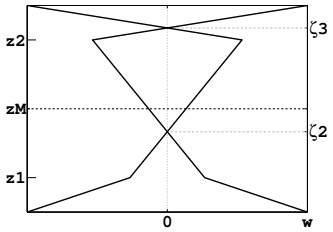
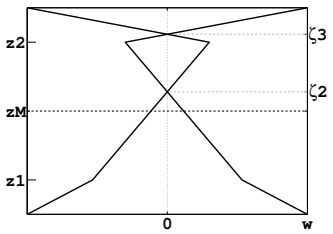
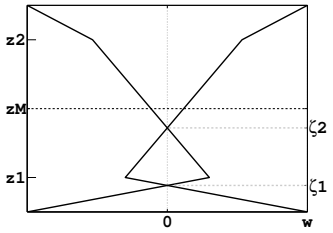
configuration	pressure force
[1] convergent open 	$F_1 = \ell P_s \left[d_1 \left(1 - \frac{a_2^2}{a_0 a_1} \right) + \frac{d_2}{2} \left(1 - \frac{a_2^2}{a_M a_1} \right) \right]$ $F_2 = \ell P_s \frac{d_2}{2} \left(1 - \frac{a_2^2}{a_M a_2} \right)$
[2] divergent open 	$F_1 = \ell P_s d_1 \left(1 - \frac{a_1^2}{a_0 a_1} \right)$ $F_2 = 0$
[3a] closed $a_1 > 0,\ a_1 \leq \lvert a_2 \rvert$ 	$F_1 = \ell P_s \left[d_1 + d_2 \left(\frac{a_1}{a_1 - a_2} \right) \right]$ $F_2 = 0$
[3b] closed $a_1 > 0,\ a_1 > \lvert a_2 \rvert$ 	$F_1 = \ell P_s \left(d_1 + \frac{d_2}{2} \right)$ $F_2 = \ell P_s \frac{d_2}{2} \left(\frac{a_1 + a_2}{a_1 - a_2} \right)$
[4] closed $a_1 \leq 0$ 	$F_1 = \ell P_s d_1 \left(\frac{a_0}{a_0 - a_1} \right)$ $F_2 = 0$

Table A.3: F values according to LTM configurations.

To solve this integral, because $a(z)$ is a linear function of z, we can write $a(z) = mz+q$ for $m,q \in \mathbb{R}$. Therefore:

$$\begin{aligned}
F_X(z',z'') &= \ell P_s(z''-z') - \ell P_s a_{min}^2 \int_{z'}^{z''} \frac{1}{(mz+q)^2}\,\mathrm{d}z \\
&= \ell P_s(z''-z') - \ell P_s \frac{a_{min}^2}{m}\left[-\frac{1}{mz+q}\right]_{z'}^{z''} \\
&= \ell P_s(z''-z') - \ell P_s \frac{a_{min}^2}{m}\left(-\frac{1}{mz''+q}+\frac{1}{mz'+q}\right) \\
&= \ell P_s(z''-z') - \ell P_s a_{min}^2\left(\frac{-z'+z''}{(mz''+q)(mz'+q)}\right) \\
&= \ell P_s(z''-z') - \ell P_s(z''-z')\left(\frac{a_{min}^2}{a(z'')a(z')}\right) \\
&= \ell P_s(z''-z')\left(1-\frac{a_{min}^2}{a(z')a(z'')}\right).
\end{aligned}$$

A.2.5 Collision forces

Traditional two-mass models do not require the calculation of contact area, because the projected area is rectangular and there is no gradation in opening and closing (Sciamarella and D'Alessandro, 2004). We define a collision force that is consistent with eq. 3.9 on p. 49 and admits a gradual variation of contact area in time. First, we remark that each collision force C_i is zero if:

$$C_i = 0 \Leftrightarrow a_i \geq 0 \text{ AND } a_M \geq 0, \quad i \in \{1,2\}$$

If the collision force C_i is nonzero, we define respectively:

$$C_1 = \frac{1}{\delta(0,d_M)} \int_0^{d_M} -\Theta(-a(z))a(z)\frac{c_i}{2\ell}\,\mathrm{d}z \tag{A.9}$$

$$C_2 = \frac{1}{\delta(d_M,d_3)} \int_{d_M}^{d_3} -\Theta(-a(z))a(z)\frac{c_i}{2\ell}\,\mathrm{d}z \tag{A.10}$$

where $\delta(z',z'')$ is the distance (on the line $x=0$) between z' and z'' along which $a(z)<0$. For simplicity, in the following we will denote $\delta(0,d_M)$ as L_1 and $\delta(d_M,d_3)$ as L_2 (see Fig. A.4).

We chose to normalize C_1 and C_2 using distance L_i in order to keep c_i of the dimension of g/ms^2 and to obtain a generalization of the previous collision force. The latter is found comparing it with the degenerate case $d_1 = 0$. For the implementation of C_i, we take into account that the integrals in eq. A.9 and A.10 are proportional either to triangular or trapezoidal areas of colliding masses: first, we observe that

$$C_i = \frac{1}{L_i}\int_{L_i} a(z)\frac{c_i}{2\ell}\,\mathrm{d}z = \frac{c_i}{2\ell L_i}\int_{L_i} a(z)\,\mathrm{d}z = \frac{c_i}{L_i}\int_{L_i} x(z)\,\mathrm{d}z.$$

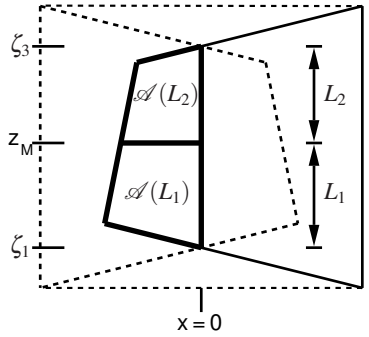


Figure A.4: The quantities L_i and $\mathscr{A}(L_i)$.

Then, defining $\mathscr{A}(L_i) = \int_{L_i} x(z)\,\mathrm{d}z$, we obtain $C_i = \frac{c_i}{L_i}\mathscr{A}(L_i)$. This (only apparent) complication is justified by the fact that the calculation of $\mathscr{A}(L_i)$ is much easier because involves simple geometry: in fact, $\mathscr{A}(L_i)$ is the colliding area on the frontal plane xz (see Fig. A.4), which can be either a triangle, or a trapezoid (see Table A.4).

If $\mathscr{A}(L_i)$ is the area of a triangle, denoting as h the height of the triangle along the x direction and z the relative ordinate (see e.g. $\mathscr{A}(L_1)$ in Fig. A.5a), we have:

$$\mathscr{A}(L_i) = \frac{hL_i}{2} \Rightarrow C_i = -\frac{c_i}{L_i}\mathscr{A}(L_i) = -\frac{c_i h}{2}$$

Because $h = \frac{a(z)}{2\ell}$, we can write:

$$C_i = -\frac{c_i a(z)}{4\ell}$$

According to the possible syringeal configurations, we have: $a(z) = a_1$ (C_1 case [a] in Table A.4), $a(z) = a_2$ (C_2 case [e]) or $a(z) = a_M$ (C_1 case [d], C_2 case [b]).

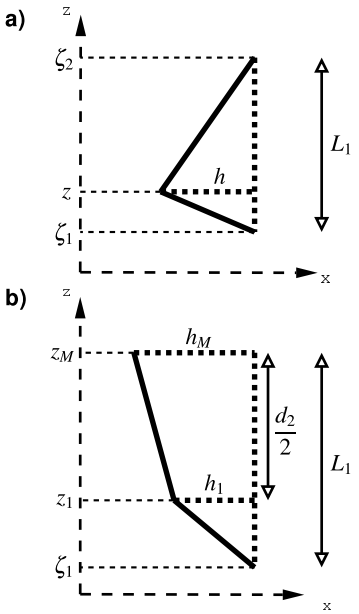


Figure A.5: $\mathscr{A}(L_1)$: triangular case (a) and trapezoidal case (b).

If $\mathscr{A}(L_i)$ is the area of a trapezoid, from simple geometry we obtain (see e.g. $\mathscr{A}(L_1)$ in Fig. A.5b):

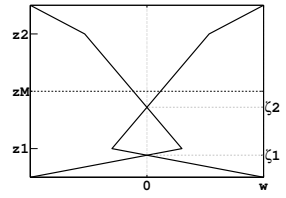
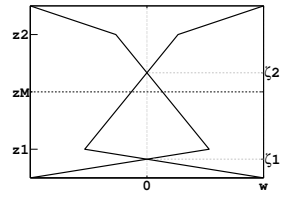
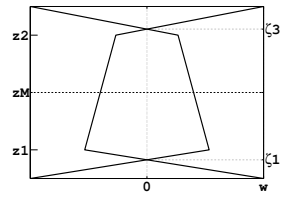
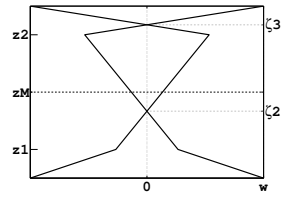
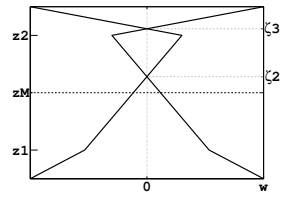
$$\mathscr{A}(L_i) = \frac{\frac{d_2}{2}(h_i + h_M)}{2} + \frac{(L_i - \frac{d_2}{2})h_i}{2} = \frac{h_i L_i}{2} + \frac{h_M d_2}{4}$$

Because $h_i = \frac{a_i}{2\ell}$ ($i \in \{1,2\}$) and $h_M = \frac{a_M}{2\ell}$, we can write:

$$\mathscr{A}(L_i) = \frac{a_i L_i}{4\ell} + \frac{a_M d_2}{8\ell}$$

and consequently:

$$C_i = -\frac{c_i}{L_i}\mathscr{A}(L_i) = -\frac{c_i}{4\ell}\left(a_i + \frac{a_M d_2}{2L_i}\right)$$

configuration	$\mathbf{C_1}$	$\mathbf{C_2}$
[a] 	$-\frac{c_1}{4\ell}a_1$	0
[b] 	$-\frac{c_1}{4\ell}\left(a_1+\frac{a_M d_2}{2(z_M-\zeta_1)}\right)$ (†)	$-\frac{c_2}{4\ell}a_M$
[c] 	$-\frac{c_1}{4\ell}\left(a_1+\frac{a_M d_2}{2(z_M-\zeta_1)}\right)$ (†)	$-\frac{c_2}{4\ell}\left(a_2+\frac{a_M d_2}{2(\zeta_3-z_M)}\right)$ (‡)
[d] 	$-\frac{c_1}{4\ell}a_M$	$-\frac{c_2}{4\ell}\left(a_2+\frac{a_M d_2}{2(\zeta_3-z_M)}\right)$ (‡)
[e] 	0	$-\frac{c_2}{4\ell}a_2$

(†) $= -\frac{c_1}{4\ell}\left[a_1+\frac{a_M d_2(a_0-a_1)}{d_2(a_0-a_1)-2a_1 d_1}\right]$

(‡) $= -\frac{c_2}{4\ell}\left[a_2+\frac{a_M d_2(a_0-a_2)}{d_2(a_0-a_2)-2a_2 d_3}\right]$

Table A.4: C values according to LTM configurations.

where $L_1 = z_M - \zeta_1$ and $L_2 = \zeta_3 - z_M$ (see, e.g., cases [c] for both C_1 and C_2 in Table A.4).

B

Bifurcation analysis and inspiration

In a typical ring dove coo (Fig. B.1), more interesting dynamics are exhibited during inspiration (or inhalation) and superfast muscle activity which controls the fast but gradual frequency modulations (trills) of ring dove coo (Elemans et al., 2004). We want to investigate the trapezoidal model to examine possible relationships between changes in model parameters and the type of vibrations.

The interpretation of the model parameters and their relationships to the actual biological forces is not a trivial task, and requires experimental validation (Mindlin et al., 2003). In Chapter 4 and Appendix C, we address this issue by integrating the simple trapezoidal model with experimental data of a standard coo to study the frequency modulation of trills. In songbirds, a similar approach has been employed by Mindlin et al. (2003). The mechanisms underlying inspiration, on the other hand, are poorly investigated and they presumably require more detailed measurements and a higher-order model with a more detailed description of the physiology. Therefore, we address here basic questions: Which parameters of the trapezoidal model could be involved during inhalation? Which parameters changes lead to changes in the radiated sound spectrum that are qualitatively comparable to the more rich harmonic spectra observed during inspiration?

We have already seen (Chapter 3) that the geometry and the rest position of the syrinx can influence the harmonic spectra drastically, and that the trapezoidal model can lead to more pronounced harmonics of the radiated sound pressure, due to collision of the masses. The spectrogram is qualitatively comparable to recorded inspiration in a typical coo of the ring dove (*Streptopelia risoria*).

The trapezoidal model described in Chapter 3 is assumed to be symmetrical between

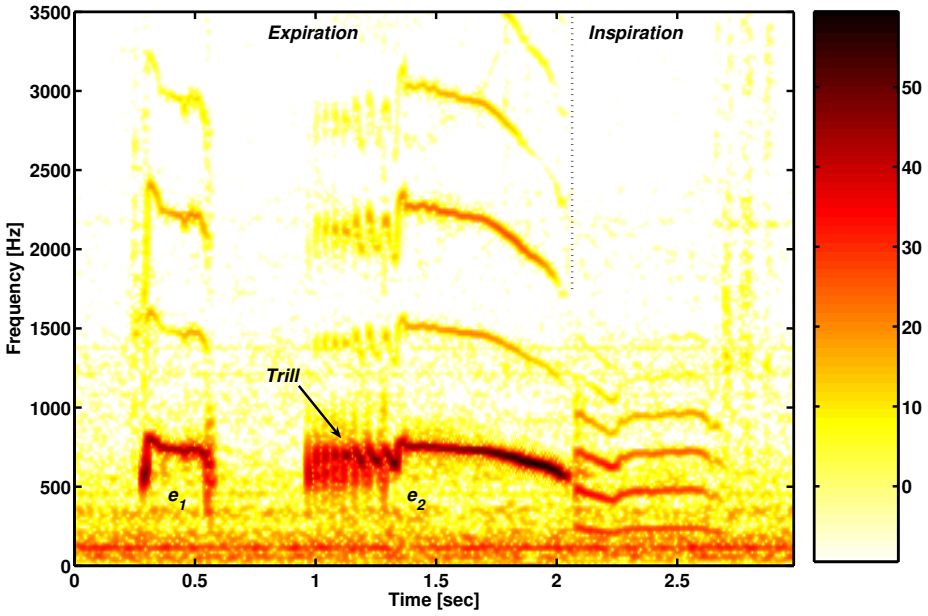


Figure B.1: The spectrogram of a typical coo of the ring dove (*Streptopelia risoria*) consists of two sounds elements e1 and e2 (Gaunt et al., 1982; Beckers et al., 2003b). The first part of e2 is amplitude modulated, which gives rise to a trill-like, rolling quality (Beckers et al., 2003b).

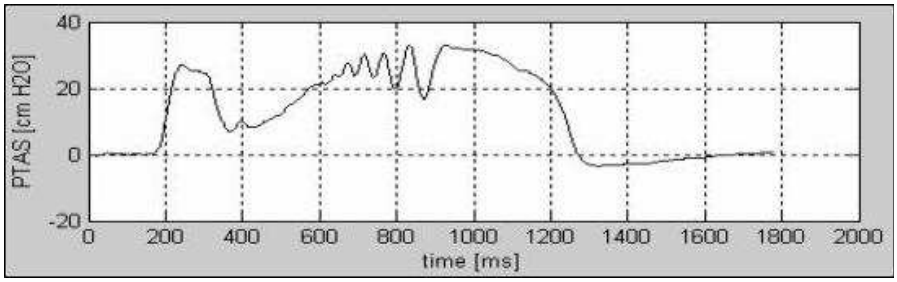


Figure B.2: Recorded P_{PTAS} (measured by C. P. H. Elemans).

upper and lower mass. We implement in this section possible asymmetries in the geometry of the model in order to investigate potential mechanisms during inspiration which can lead to more pronounced harmonics and oscillations at relatively low pressure (see Fig. B.2: inspiration starts at 1260 msec. The pressure gradient $|\Delta P|$ acting on the labia, assuming $P_{\text{sup}} = 0$, is in the order of 3-4 $\text{cm}\,\text{H}_2\text{O}$). As pointed out in Chapter 3, there is almost no information available on the prephonatory shape of vibrating tissues in avian phonation studies. Therefore, we will explore changes in the syringeal width and in the syringeal shape. The former is represented by the parameter a_{0i} (assuming for simplicity $a_{01} = a_{02}$), the latter by the ratio $\frac{a_{02}}{a_{01}}$, which induces an asymmetry between lower and upper rest areas. Furthermore, the shape of the syrinx allow us to slightly alter the length of the first plate d_1 and study asymmetries between lower and upper plate.

The bifurcation diagrams presented here were obtained by means of the software XPPAUTO, spectrograms and simulations by means of code and Graphical User Inter-

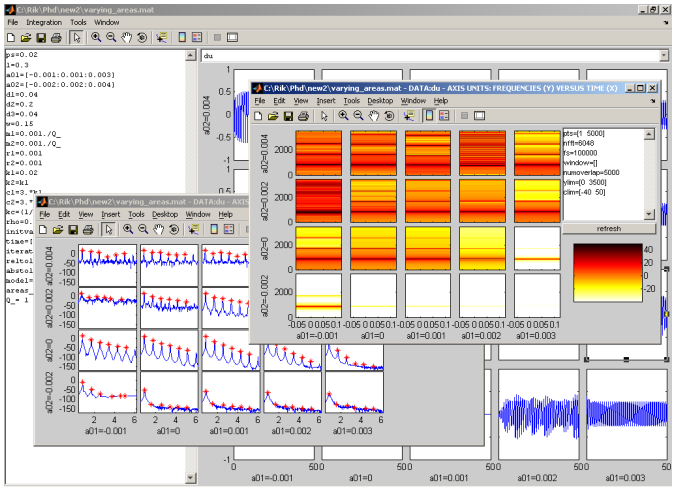


Figure B.3: The GUI of the MATLAB code employed in the simulations.

faces (GUIs) implemented in MATLAB (Fig. B.3).

B.1 Bifurcation diagrams

Prephonatory shape and syrinx geometry are analyzed in this section. We relate these results to possible mechanisms underlying inspiration. Bifurcation diagrams of the variables x_1 and x_2 versus the pressure P_s with the standard set of parameters (see Table 3.2 on p. 54) are shown in Fig. B.4b, B.6b and B.7b. All units are given in grams, centimeters, milliseconds and their relative combinations. With the standard set of parameters we found coexistence of a small and a large limit cycle (we will denote them as Λ_1 and Λ_2, respectively) in the range 425-535 Pa[1] (see Chapter 3 and Fig. B.4b).

For a convergent syrinx (or convergent prephonatory shape, i.e. $a_{01} > a_{02}$, see Fig. B.4a) we observe an increased onset and a longer duration of Λ_2. However, the bifurcation diagrams are qualitatively comparable to the rectangular case (Fig. B.4a): Λ_1 is collision-free, transitions from Λ_2 (with collisions) to Λ_1 are observed due to a loss of stability of Λ_2.

A divergent syrinx, on the other hand, leads to more instabilities (Fig. B.4c): the software XPPAUTO detected bifurcation points (BP) and period doubling bifurcation PD (for details, see Ermentrout, 2001) which do not appear in the case of a rectangular and

[1] We remark that $100\ \text{Pa} = 0.001 \frac{\text{g}}{\text{cm ms}^2} \simeq 1\ \text{cm}\,\text{H}_2\text{O}$.

convergent glottis. Moreover, the Λ_2 seems to loose the stability at about 800 Pa (0.008 $\frac{\text{g}}{\text{cm ms}^2}$, see Fig. B.4b) but no stable branch was detected by XPPAUTO. Oscillations run for increasing P_s values, however, seem to confirm a region were chaotic oscillations occur (see Fig. B.5). At $P_s = 0.012$ $\frac{\text{g}}{\text{cm ms}^2}$ (12 cmH_2O), a jump to a qualitatively different limit cycle is observed.

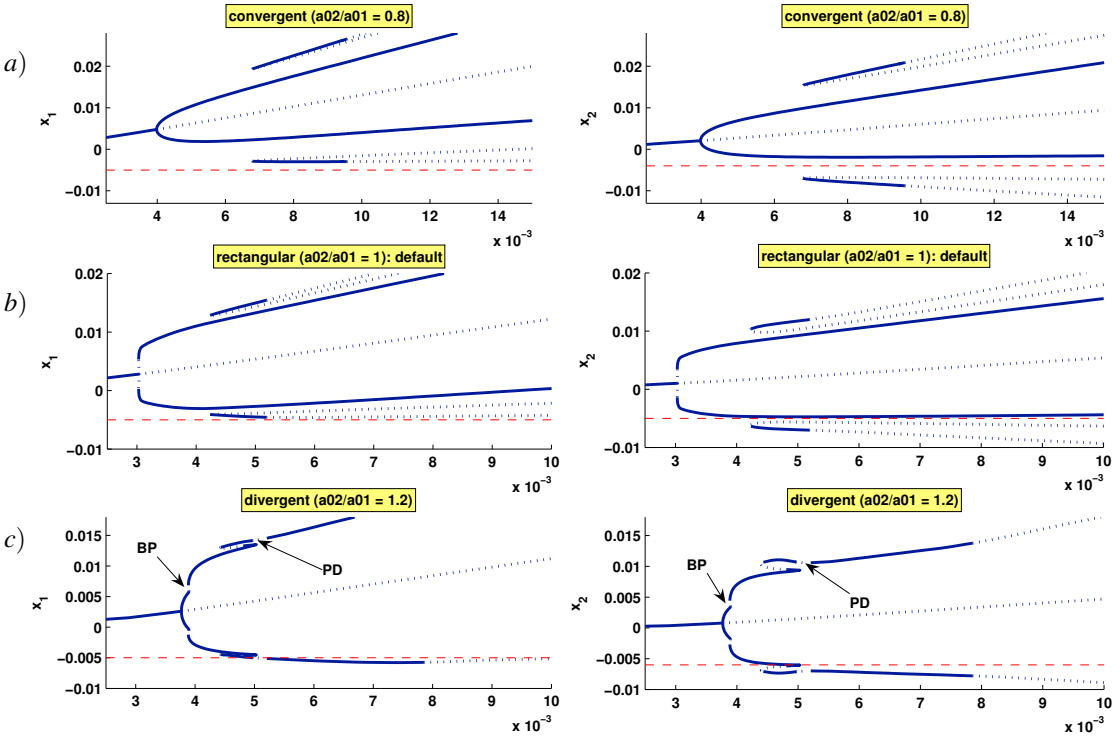


Figure B.4: Bifurcation diagrams of the variables x_1 and x_2 versus P_s (in $\frac{\text{g}}{\text{cm ms}^2}$) for (a) divergent, (b) rectangular and (c) divergent syrinx. Dotted lines represent unstable limit cycles or fixed points, solid lines stable limit cycles or fixed points. Values of x_i, $i \in \{1,2\}$ below the dashed horizontal line (in red) represents collision of the masses m_i, $i \in \{1,2\}$.

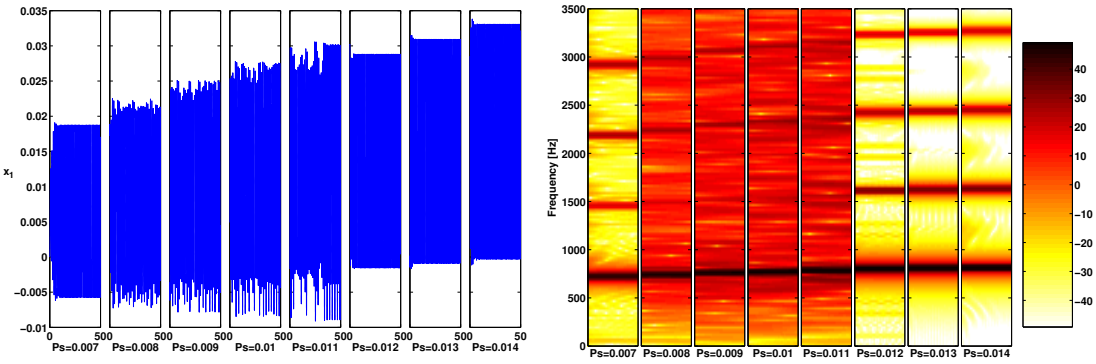


Figure B.5: Oscillations for increasing pressure and a divergent glottis (see Fig. B.4c) confirm regions were chaotic oscillations occur. Left panel: x_1 versus P_s and spectrogram of the relative $\dot{U}$ (right panel). At approximately 0.012 $\frac{\text{g}}{\text{cm ms}^2}$, small increase of subsyringeal pressure seems to induce a sudden jump to a qualitatively different (periodic) oscillation.

Interestingly, the situation just described for changes in the prephonatory shape of the syrinx is similar if we vary the first plate elongation (Fig. B.6). An increase of the first plate length ($d_1 = 0.06$ cm, see Fig. B.6a) leads to bifurcation diagrams qualitatively similar to the case of the convergent syrinx, whereas a reduction of the first plate length ($d_1 = 0.02$ cm, see Fig. B.6c) leads to more instabilities such as period doublings. However, no bifurcation points were detected by XPPAUTO (compare Fig. B.4c).

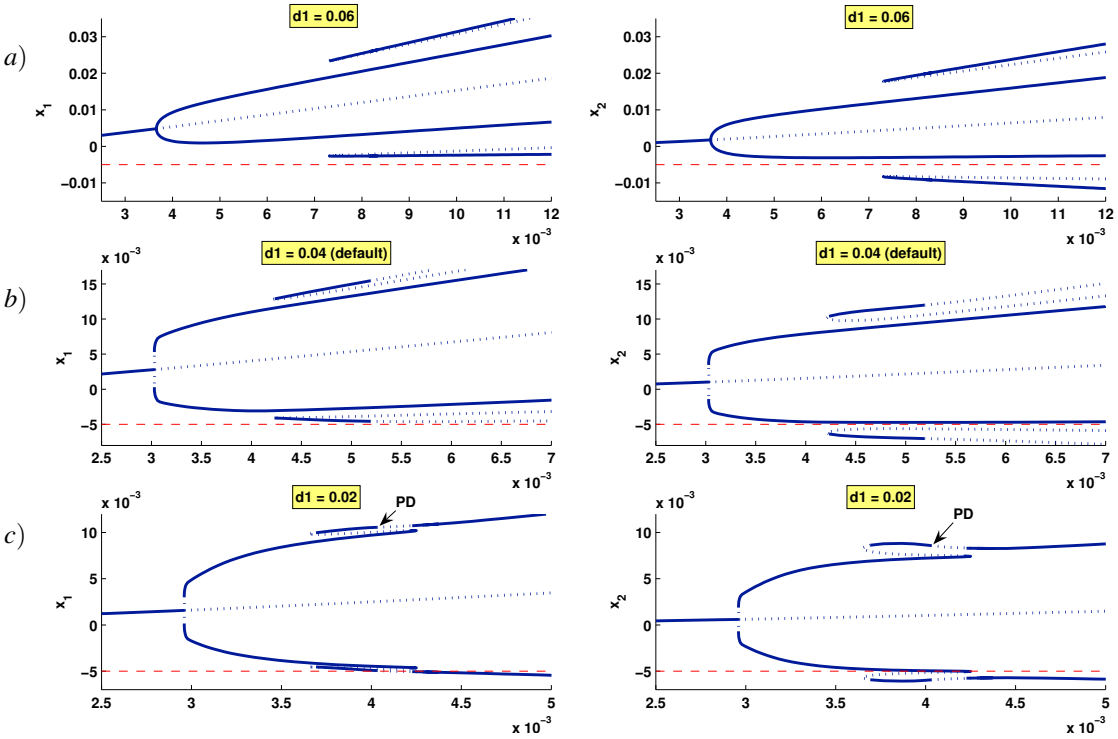


Figure B.6: Bifurcation diagrams of the variables x_1 and x_2 versus P_s (in $\frac{\text{g}}{\text{cm ms}^2}$) for (a) $d_1 = 0.06$ cm, (b) $d_1 = 0.04$ cm, (c) $d_1 = 0.02$ cm. Dotted lines represent unstable limit cycles or fixed points, solid lines stable limit cycles or fixed points. Values of x_i, $i \in \{1,2\}$ below the dashed horizontal line (in red) represent collision of the masses m_i, $i \in \{1,2\}$.

A reduction of the rest areas does not induce qualitative changes in the bifurcation diagrams of the variable x_i versus P_s (see Fig. B.7), but the onset of oscillations decreases according to a_{0i}.

B.2 Simulations of inspiration

The bifurcation analysis suggests that implementation of asymmetries can lead to qualitative changes of the output spectrum: we have seen that a divergent configuration leads to more instabilities and that onset of oscillations is reduced by reducing the rest areas.

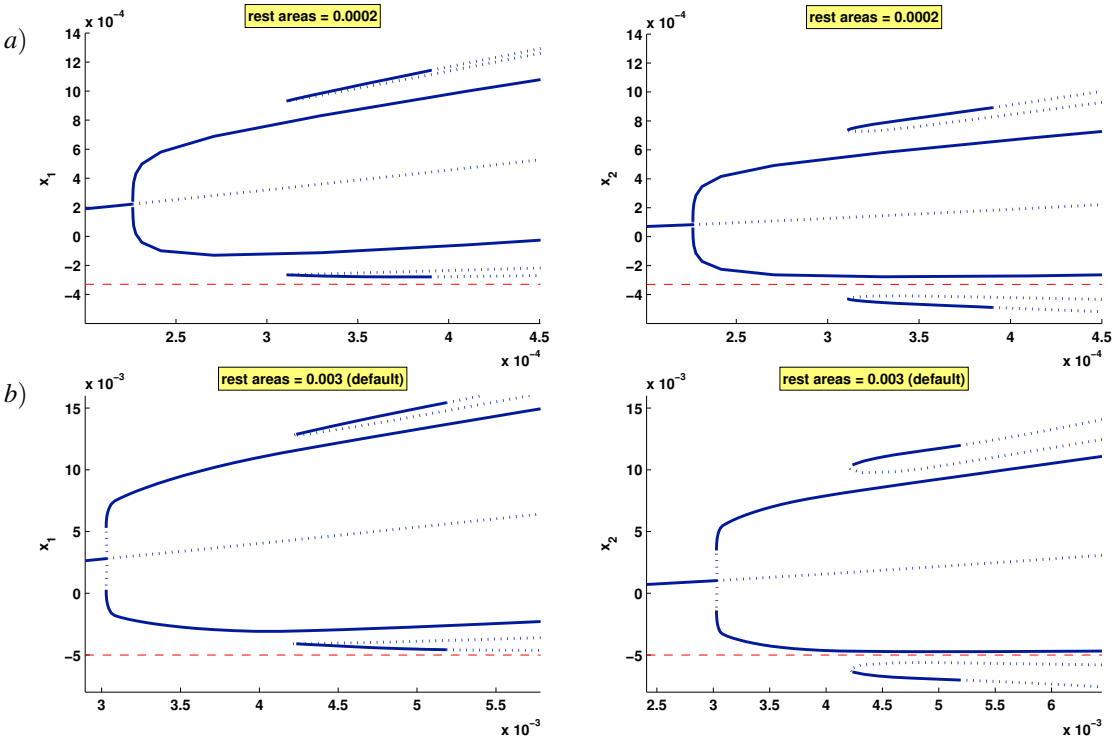


Figure B.7: Bifurcation diagrams of the variables x_1 and x_2 versus P_s (in $\frac{\text{g}}{\text{cm ms}^2}$) for (a) almost closed syrinx ($a_{01}, a_{02} \simeq 0$ cm) and (b) default rest areas values ($a_{01}, a_{02} = 0.003$ cm). Dotted lines represent unstable limit cycles or fixed points, solid lines stable limit cycles or fixed points. Values of x_i, $i \in \{1,2\}$ below the dashed horizontal line (in red) represent collision of the masses m_i, $i \in \{1,2\}$.

Before proceeding in the simulation of the inspiratory part of the coo by varying the model parameters related to rest areas and syrinx prephonatory shape, namely a_{0i} ($i \in \{1,2\}$) and $\frac{a_{02}}{a_{01}}$, we run several bifurcation diagrams of the pressure P_s versus the model parameters in order to gain further insights on possible mechanisms underlying inspiration (Fig. B.8). Moreover, Fig. B.8a and Fig. B.8h will be recalled in Appendix C when deriving the scaling functions $Q(t)$ and $a_{0i}(t)$, respectively (compare Fig. 4.4e-d on p. 76). The diagrams confirm that the onset of oscillations is reduced by reducing the rest areas (Fig. B.8h, right branch) and that the onset of oscillations is lower for a convergent prephonatory shape (Fig. B.8d, right branch) than for a divergent prephonatory shape (Fig. B.8d, left branch). Interestingly, for an almost closed syrinx, onset of oscillations is lower for a *divergent* syrinx than for a convergent syrinx. In other words, by setting $a_{01}, a_{02} \simeq 0$ as default, the bifurcation diagram of B.8d changes qualitatively (see Fig. B.9a versus Fig. B.9b-c). It is commonly known (Scherer and Titze, 1983; Titze, 1988) that a rectangular or slightly convergent configuration is optimal for the onset of oscillations in vocal folds, contrarily to divergent or strongly convergent configurations. In the trapezoidal model, *for rest areas close to zero*, a divergent configuration seems to be optimal for the onset of oscillation. However, if we assume a divergent prephonatory shape and we increase P_s, close to the Hopf bifurcation (onset of oscillation), the shape

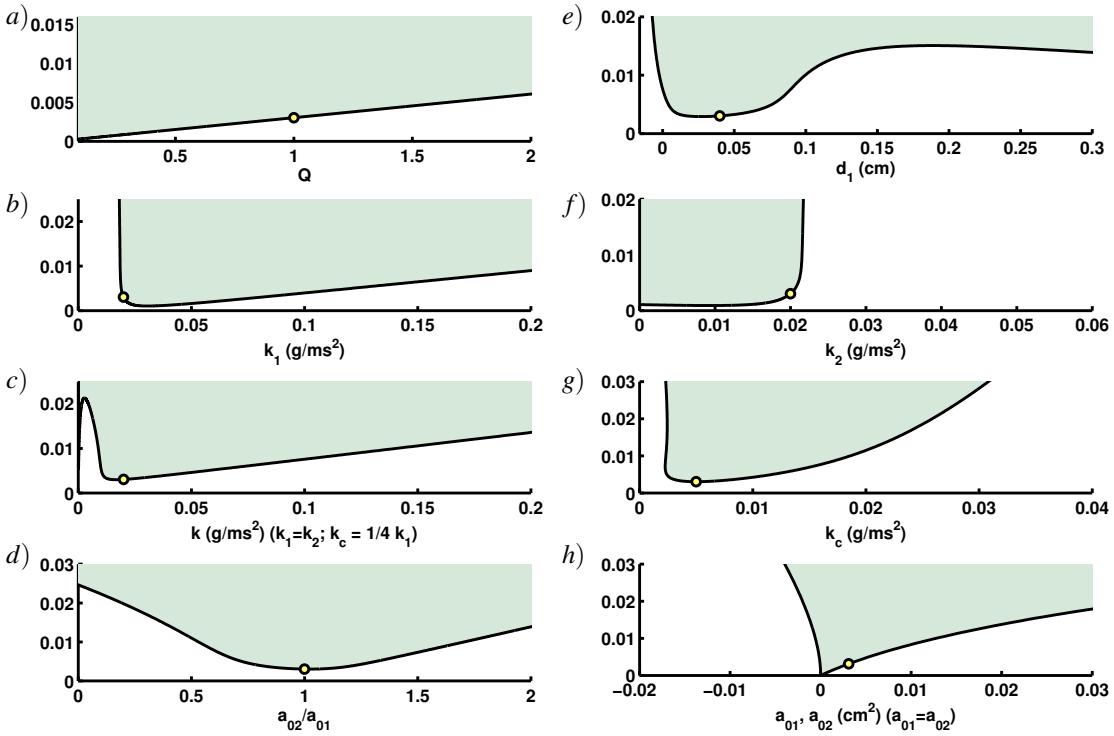


Figure B.8: Bifurcation diagrams of the trapezoidal model: Oscillation onset pressure (in $\frac{\text{g}}{\text{cm ms}^2}$) versus different parameters. The colored area above each bifurcation line denotes the region where oscillation of the masses occurs. The dot represent the default parameter value.

of the syrinx (syringeal areas a_1 and a_2) becomes rectangular or even slightly convergent (Fig. B.10a). On the other hand, a convergent prephonatory shape causes the labia to increase the convergence, therefore to move away from the optimal configuration (Fig. B.10b). More pressure is then required to trigger oscillations of the labia.

B.2.1 Simulations setup

In order to perform simulations of the inspiratory part of the coo, the expiratory part of the coo was simulated by means of time-dependent control parameters, namely $P_s(t)$, $a_{0i}(t)$ and $Q(t)$.

The subglottal pressure $P_s(t)$ was build from experimental measurements of the P_{PTAS} (see Fig. B.2) by picking some point of the pressure trace and connecting them linearly. The ripples in the P_{PTAS} (between 700 and 900 msec in Fig. B.2) were omitted for simplicity. The rest areas a_{0i} are assumed to be constant and in a slightly convergent shape ($a_{01} > a_{02}$).

The function $Q(t)$ is a time-dependent parameter which reproduces the frequency modulation (FM) of the coo by affecting masses and tension parameters (see Appendix C for details). $Q(t)$ is assumed for simplicity to follow the same time course described by $P_s(t)$. First, we scaled $P_s(t)$ to obtain a normalized function $Q_{norm}(t) \in [0,1]$ in order

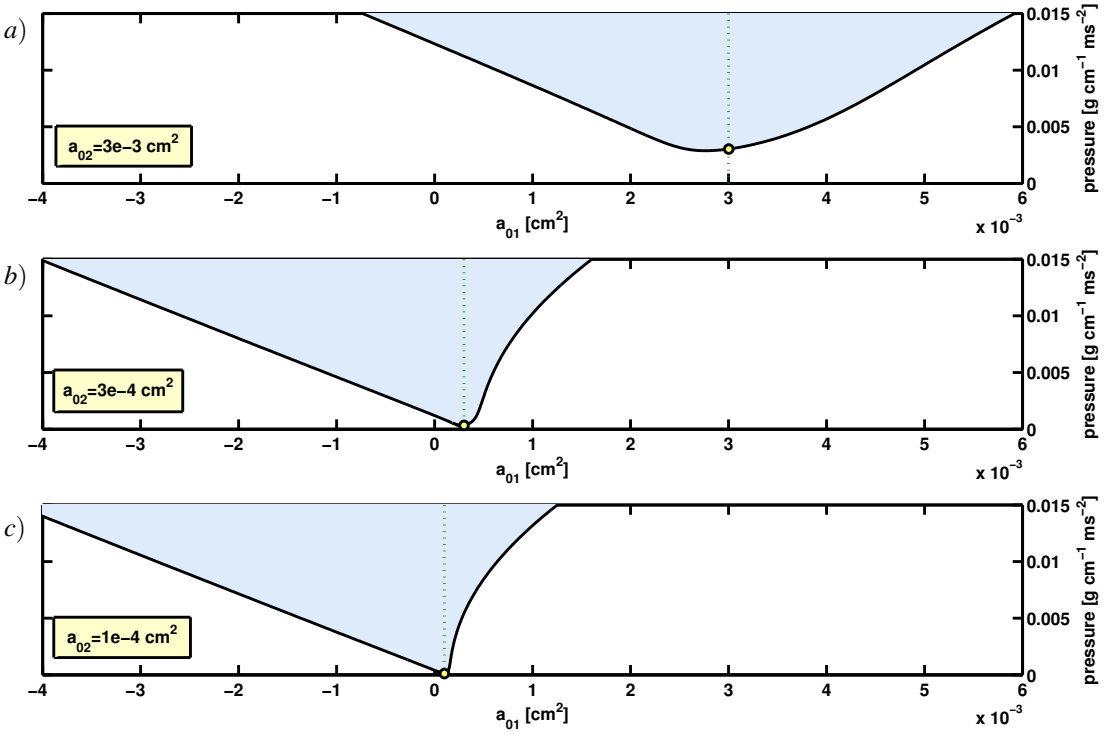


Figure B.9: Oscillation onset pressure versus a_{01}. On each panel, the shaded area represents oscillation of the masses. The dot on the bifurcation line indicates the default rest areas values of the simulation, i.e., a rectangular prephonatory shape ($a_{01} = a_{02}$). Consequently, as in Fig. B.8d, each point located at the left of the vertical dotted line represents a divergent prephonatory shape ($a_{02} > a_{01}$, cf. the default constant value of a_{02} in the bottom left frame), whereas each point located at the right represents a convergent shape ($a_{02} < a_{01}$). Decreasing the rest areas (panel a) to c)) the qualitative changes in the bifurcation lines show a lower onset for a divergent prephonatory shape.

to obtain approximately the same fundamental frequency of the recorded coo (Fig. B.1) in our simulations. Second, we built the trill as a sinusoidal function $\tau(t)$ with frequency at approximately 30 Hz (Elemans et al., 2004) and amplitude $A \simeq \frac{1}{10}$ chosen to match the frequency modulation of the trill (with this choice, FM ranges from $0.9\mathscr{F}_0$ to $1.1\mathscr{F}_0$ during trill). The function $\tau(t)$ is defined only for t values within trill onset t_1 and trill offset t_2, otherwise $\tau(t) = 0$. The numbers t_1, t_2 were estimated on the basis of Fig. B.1. In symbols $Q(t) = Q_{norm}(t) + \tau(t)$, where

$$\tau(t) = \frac{1}{10} \sin\left(2\pi \frac{30(t-t_1)}{1000}\right) \Theta(t-t_1)\Theta(t_2-t)$$

($\frac{t}{1000}$ is due to the chosen unit in milliseconds).

Two aspects justify our oversimplified assumptions: these simulations were run with the experimental data reported in Chapter 4 not yet available, and we can test the behaviour of the trapezoidal model under the effect of time dependent driving parameters. Concerning this last point, it is worth to remark that simulations of the expiratory part

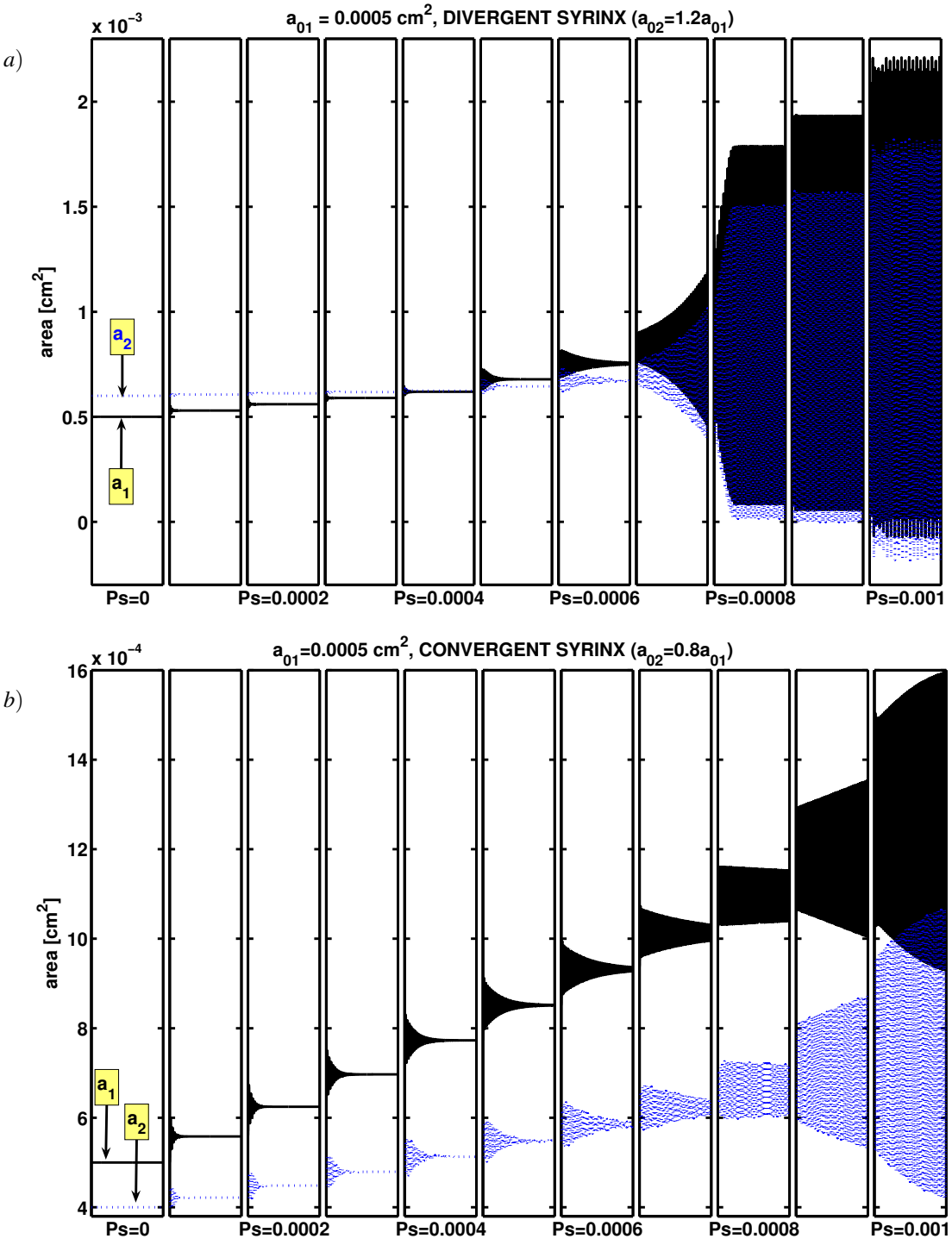


Figure B.10: Syringeal areas oscillations for narrow prephonatory shape (rest areas close to zero) and increasing P_s (x axis) in $\frac{\text{g}}{\text{cm ms}^2}$. Panel (a): slightly divergent syrinx. P_s (always acting on the first mass via the lower plate d_1) causes the oscillating areas a_1, a_2 to reach the optimal configuration for the onset of oscillations (slightly convergent) at a value $\approx 0.0007 \frac{\text{g}}{\text{cm ms}^2}$ ($\simeq 0.7$ cmH_2O). Panel (b): slightly convergent syrinx. P_s increases the "convergency" $\frac{a_{01}}{a_{02}}$ and more pressure is needed to sustain oscillation (onset at $\approx 0.0008 \frac{\text{g}}{\text{cm ms}^2}$).

of the coo resemble qualitatively a typical ring dove coo (Fig. B.11). Even with our basic assumptions, this is not a trivial result: we showed in Chapter 3 that in the trapezoidal model different parameter values can lead to distinct oscillation patterns, and small perturbations induced jumps from one attractor to the other. In general, nonlinear dynamical systems (see Section 2.2) make the principle of superposition no longer applicable, and different excitations may result in totally different time-courses. In a typical ring dove coo (and in our first simulations), register transitions to aphonia, period doublings or chaos are rare during expiration or inspiration.

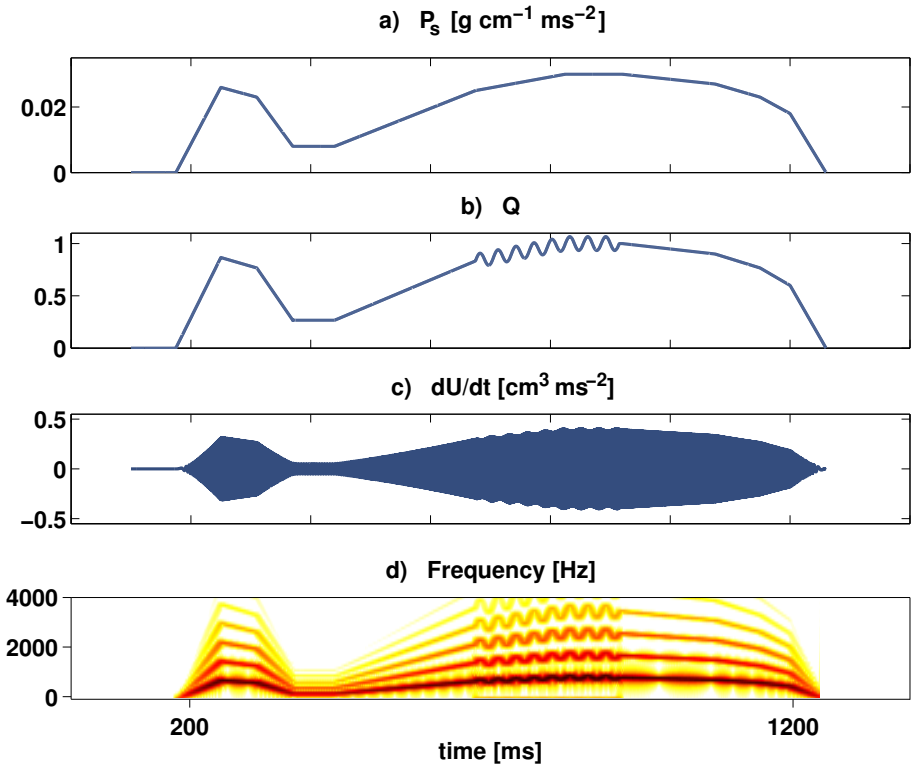


Figure B.11: First simulation of a ring dove coo. Driving parameters (a-b), (c) oscillogram of the flow derivative and (d) relative spectrogram.

B.2.2 Results

During inspiration, we assume the pressure gradient acting on the labia to be $P_s(t) = |P_{PTAS}|$, i.e., we assume $P_s = P_{\text{sup}}$ and $P_{\text{sup}} = 0$. In this "reversed" condition, the prephonatory shape is set as consequence from slightly convergent to slightly divergent. Furthermore, we adduct the labia in order to obtain oscillations at relatively low onset. Finally, the fundamental frequency is assumed to be constant in the order of $\frac{1}{3}$ of the standard $\mathscr{F}_0$ of the trapezoidal model to match approximately the fundamental frequency of the coo.

Simulations based on these assumptions are shown in Fig. B.12, where the expiratory part ends at about 1260 milliseconds, followed by the onset of inspiration: $P_s \simeq |P_{PTAS}|$ (Fig. B.12a), the fundamental frequency is controlled by $Q(t) \simeq \frac{1}{3}$ (Fig. B.12b), rest areas adduct and the prephonatory shape becomes divergent (Fig. B.12c). The resulting spectrogram is shown in Fig. B.12d. As we can see, a more rich harmonic spectrum characterizes the inspiratory part of the coo, in qualitative agreement with the recorded sounds of the ring dove coo during inhalation and consistently with the bifurcation analysis (Fig. B.4,B.6, B.7), which suggests that such spectrograms are in general obtainable for a divergent configuration and rest areas close to zero.

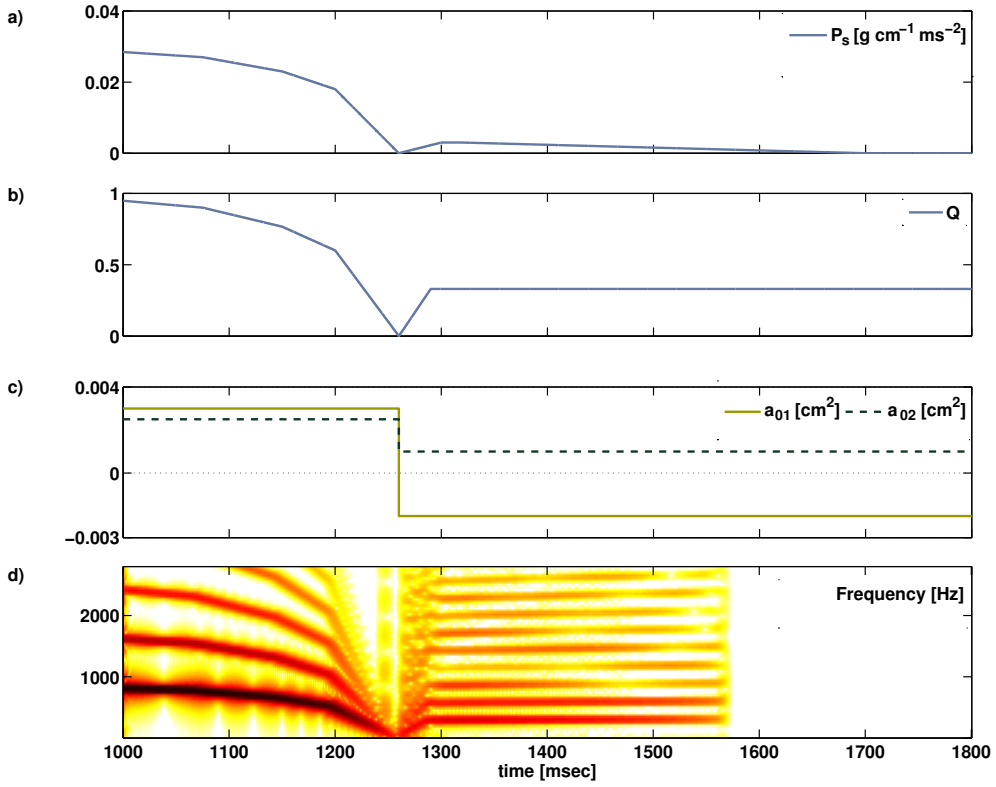


Figure B.12: Simulation of inspiration by means of time-series describing the changes of parameters in the trapezoidal model. a) $P_s(t)$, b) $Q(t)$ which controls the fundamental frequency, c) rest areas and d) spectrogram of the resulting flow derivative. At inspiration (1260 msec), rest areas are "reversed" and slightly adducted (moved towards zero).

In Fig. B.13 we run several simulations with the same time courses employed in Fig. B.12 and varying in turn rest areas a_{01} (x axis) and a_{02} (y axis). We framed the spectrograms qualitatively comparable to Fig. B.12d according to the relative energy (e.g., color) of the second and third harmonic with respect to the fundamental. These spectra are located in the region denoting a divergent prephonatory shape (e.g., on the top left surface with respect to the diagonal dashed line) and rest areas close to zero (e.g., close to the diagonal dashed blue line). Oscillations are also set on for an open rectangular syrinx and rest areas close to zero (spectrograms along the diagonal dashed

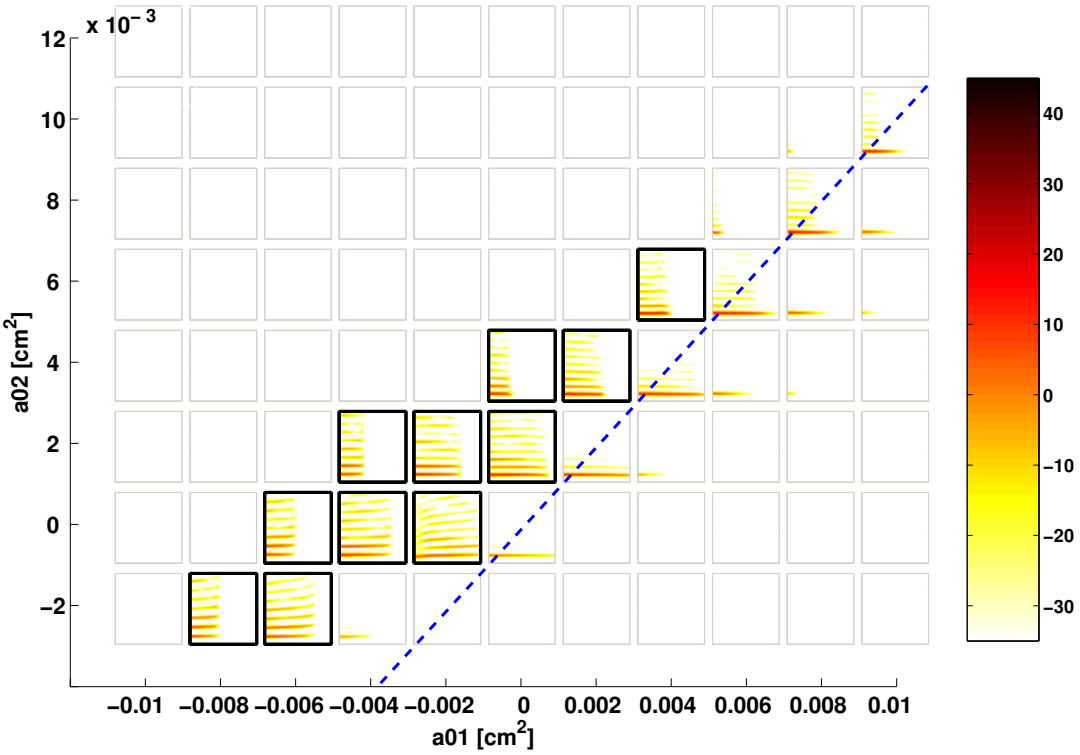


Figure B.13: Simulations of the inspiratory part of the coo for varying rest areas (compare Fig. B.12d after approx. 1260 milliseconds) produce spectrograms showing strong harmonics (dark frame) in the region denoting a divergent prephonatory shape (i.e., on the top left surface with respect to the diagonal dashed blue line) and rest areas close to zero (i.e., close to the diagonal dashed blue line).

blue line), but, probably due to the syrinx geometry, no collision occur and therefore the energy at the fundamental is relatively stronger than the harmonics.

It is likely that more mechanisms than those suggested here are involved during inspiration of a ring dove coo. The second harmonic of inspiration seems to follow the fundamental frequency of the expiratory part (Fig. B.1), suggesting the appearance of period doublings. Period doublings appeared also in several simulations with the trapezoidal model (see Fig. B.4, B.6 and Chapter 3). However, we cannot find a quantitative agreement because of the lack of experimental data and the oversimplified physiology of the trapezoidal model. Moreover, the scaling of the frequency during inspiration is not supported by experimental validation, but is set in order to match the fundamental frequency of the coo. However, we have confirmed that the geometry and rest position of the syrinx can influence the radiated sound spectrum drastically and that a divergent prephonatory shape could underlie the more rich harmonic spectrum of the inspiratory part of the ring dove coo. Adduction of the labia could be involved in this process to allow oscillations at relatively low pressures.

C

Implementation of time dependent parameters in the trapezoidal model

In Chapter 4, we argue that the syringeal muscles, bronchial pressure and the transmural pressure affect the geometry and tension of the LTM. We hypothesize that

① activity by the TL muscle (P_{TL}) modulates syringeal aperture and

② the sum of the transmural pressure and muscle stress ($P_{sum} = P_t + P_{\mathrm{TL}}$) modulates LTM tension.

We implemented the above hypotheses in our model:

① The modulation of rest areas (i.e. syringeal aperture) described by the function $a_{0i}(t)$ is a linear function of P_{TL}, the stress generated by the tracheolateralis muscle, TL (Fig. 4.3d).

② The tension modulation, described by the function $Q(t)$ is a linear function of $P_{sum}(t)$.

We describe in this section the derivation of the functions $a_{0i}(t)$ and $Q(t)$.

C.1 Scaling functions

C.1.1 Rest areas

As discussed in Chapter 4, the measured values of the stress due to TL (function $P_{\mathrm{TL}}(t)$) are in the range $[\min P_{\mathrm{TL}}, \max P_{\mathrm{TL}}]$ = [0,20] kN/m^2. Simulations of the model indicate that the corresponding range at rest areas a_{0i} ($[\min a_{0i}, \max a_{0i}]$) is between -3 and 1 mm^2 (-0.03 and 0.01 cm^2 in the default model units, respectively). To obtain the function $a_{0i}(t)$ we used the scaling function

$$\frac{a_{0i}(t) - \min a_{0i}}{\max a_{0i} - \min a_{0i}} = \frac{P_{\mathrm{TL}}(t) - \min P_{\mathrm{TL}}}{\max P_{\mathrm{TL}} - \min P_{\mathrm{TL}}} \tag{C.1}$$

Eq. C.1 is the linear transformation which scales P_{TL} (t) between $\min a_{0i}$ and $\max a_{0i}$. Geometrically, Eq. C.1 describes the line passing for the points $(\min P_{\mathrm{TL}}, \min a_{0i})$ and $(\max P_{\mathrm{TL}}, \max a_{0i})$. Therefore, because $[\min P_{\mathrm{TL}}, \max P_{\mathrm{TL}}] \simeq [0, 20]$, denoting $\Delta a = \max a_{0i} - \min a_{0i}$, eq. C.1 leads to:

$$a_{0i}(t) \simeq P_{\mathrm{TL}}(t) \left(\frac{\Delta a}{20} \right) + \min a_{0i} \tag{C.2}$$

In order to choose numeric values of the bounds $\min a_{0i}$ and $\max a_{0i}$, we looked at the function P_{TL} in relation to the sound off- and on-set of the coo spectrogram (Fig. C.1a). For instance, assuming a constant average pressure P_s during coo of approximately 15 cm H_2O (0.005 $\frac{\mathrm{g}}{\mathrm{cm\ ms^2}}$) based on air sac pressure measurements (see Chapter 4), we could translate the bounds $\min P_{\mathrm{TL}} = 0$, $\max P_{\mathrm{TL}} = 20$ into $\min a_{0i}$ and $\max a_{0i}$ by means of the bifurcation diagram of the onset pressure versus rest areas (Fig. C.1b), following a general rule of thumb dictated by consistency reasons:

- No oscillation (sound offset) at each local minima of P_{TL}: $\min a_{0i}$ must be a value on the x axis (Fig. C.1b) whose corresponding onset (on the y axis) is higher than P_s
- Oscillation (sound onset) at each local maxima of P_{TL}: $\max a_{0i}$ must be a value on the x axis (Fig. C.1b) whose corresponding onset (on the y axis) is lower than P_s
- Preserve proportions: Sound offset happens at approximately $\frac{3}{4}$ of the maximal P_{TL} stress (see Fig. C.1a). Therefore, we expect the onset of oscillations to correspond at approximately $\frac{3}{4}$ of the maximal $a_{0i}(t)$ elongation (Fig. C.1b).

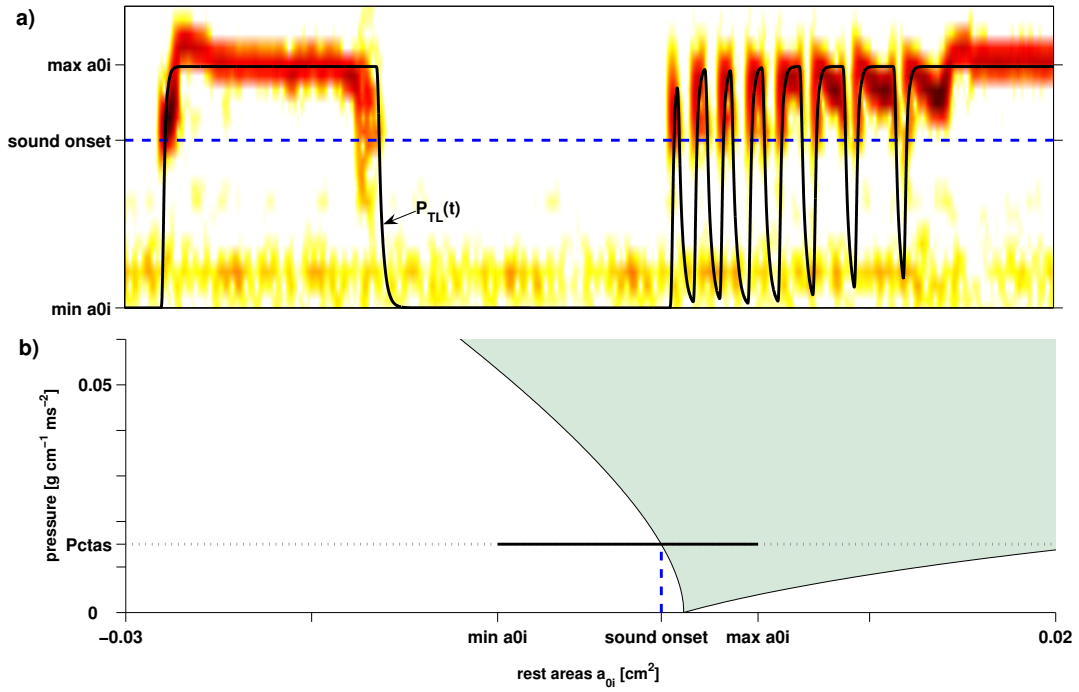


Figure C.1: Scaling functions of the rest areas $a_{0i}(t)$ controlled by TL activity. (a) TL activity (P_{TL}, solid line) is compared to the corresponding sound spectrogram. On- and offset of oscillations are estimated and (b) translated into possible $\min a_{0i}$ and $\max a_{0i}$ values according to the bifurcation diagram describing the onset of oscillations for varying rest areas a_{0i}.

However, because P_s is not constant during trill and the mechanics underlying a coo are quite complex, we run several simulations with varying rest areas bounds $\min a_{0i}$ and $\max a_{0i}$.

In Fig. C.2a-i, the function $a_{0i}(t)$ has extremes $\min a_{0i}$ and $\max a_{0i}$ which are shown in the first column (the shaded area denotes the range of rest areas values according to the bifurcation line describing the sound onset). Simulations run with $a_{0i}(t)$ produce oscillating areas a_1 and a_2, flow derivative $\dot{U}$ and relative spectrogram which are shown in the second, third and fourth column, respectively. As we can see, the gate of trill is not trivial. For small values of Δa (Fig. C.2a-d) some of the trill elements are not reproduced, as well as for larger values of Δa (Fig. C.2f-h). Trill is successfully reproduced if we increase Δa and decrease $\min a_{0i}$ (Fig. C.2i), or if we increase both $\min a_{0i}$ and $\max a_{0i}$ (Fig. C.2e). In the latter case, however:

① the scaling bounds $[\min a_{0i}, \max a_{0i}] = [-0.002, 0.01]$ are not consistent with former reasoning (preserving proportions)

② as discussed in Chapter 4, simultaneous pressure and flow recordings suggest that the syrinx is mostly closed in between sound elements during trill (Gaunt et al., 1982; Beckers et al., 2003b). This is not the case here: because the oscillating

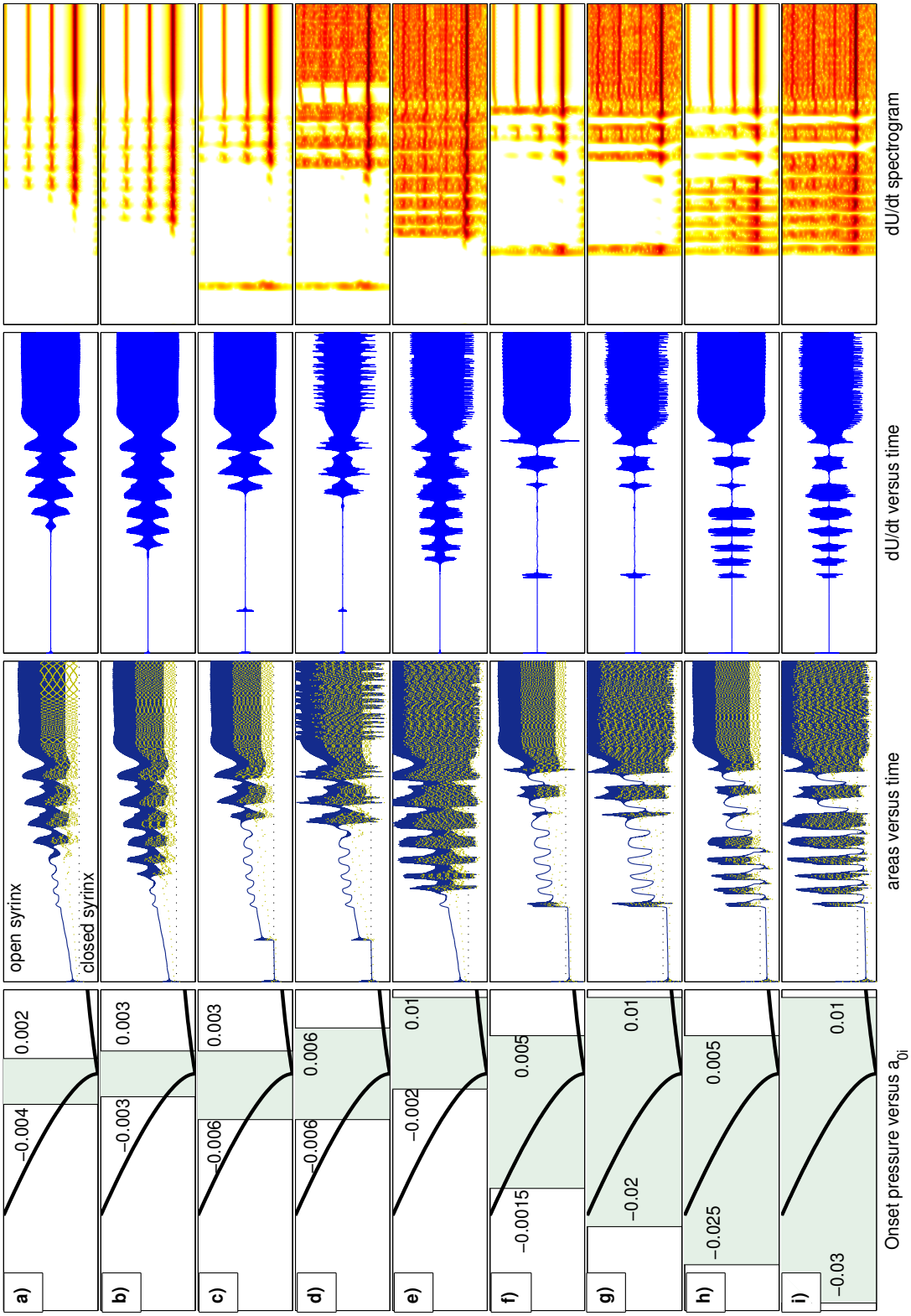


Figure C.2: The scaling functions $a_{0i}(t)$ implemented with different ranges $[\min a_{0i},\ \max a_{0i}]$ (shaded area in figures a) to i), first column) produce oscillating areas a_1 and a_2, flow derivative $\dot{U}$ and relative spectrogram which are shown in the second, third and fourth column, respectively (In figures a) to i), first column, superimposed on the shaded area is the bifurcation diagram of the onset pressure versus a_{0i}).

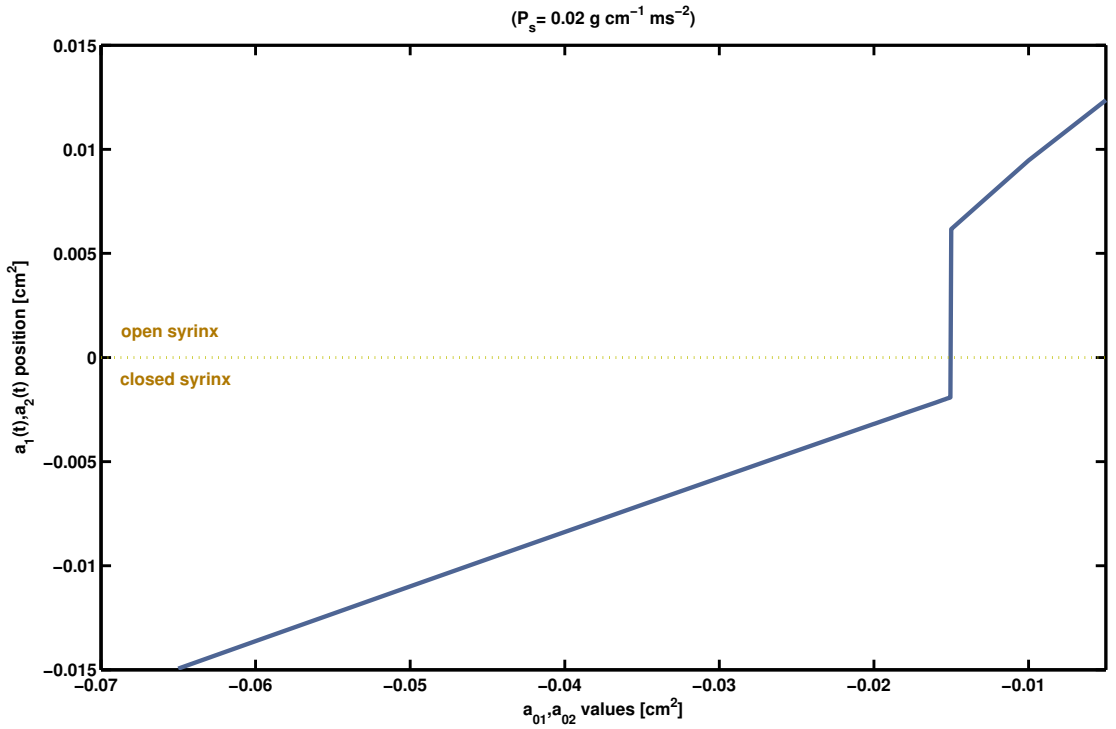


Figure C.3: Oscillating area values (y axis) for given prephonatory glottal width (x axis). The line represents stable fixed points (no oscillations). Onset of oscillation starts at rest areas values above approximately -0.005 cm^2.

area are greater than zero (Fig. C.2e, second column), the syrinx does not close in between each trill element.

As we can see in Fig. C.3 the lower the rest areas, the lower the resulting value of a_i, and consequently more pressure is needed to open the syrinx (we remark that P_s is not constant during the coo. In Fig. C.3 we show simulation with a constant pressure of 0.02 $\frac{\text{g}}{\text{cm ms}^2}$ to approximate the values of P_s in the silent periods between each trill element). Concluding, experimental evidence and simulations suggest that a consistent range is $[\min a_{0i}, \max a_{0i}] = [-0.03, 0.01]$ cm^2, as depicted in Fig. C.2i.

C.1.2 Frequency modulation

In Section C.1.1, we introduce a scaling factor to transform the range of P_{TL} values linearly to the range of rest areas. In a similar way we rescale the range of P_{sum} to Q values between 0.8 ($\min Q$) and 1.2 ($\max Q$) to match the frequency range of the coo (see Fig. 4.4e on p. 76).

We implemented tension modulation of the LTM by means of parameter $Q(t)$ which re-defines the masses and the stiffness of the LTM as a function of time: $m(t) = \frac{m}{Q(t)}$, $k(t) = kQ(t)$, where mass m and stiffness k are the default values (Table 4.2). Several implementations of Q have been used in Ishizaka and Flanagan (1972) and Steinecke

Description	Symbol	Old value	New value (increase in %)
mass (g)	m $(m_1 = m_2)$	$1.0 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$ (+70%)
stiffness ($\frac{g}{ms^2}$)	k $(k_1 = k_2)$	$20.0 \cdot 10^{-3}$	$22.0 \cdot 10^{-3}$ (+10%)
damping constant ($\frac{g}{ms}$)	r	$1.0 \cdot 10^{-3}$	$1.2 \cdot 10^{-3}$ (+20%)
coupling constant ($\frac{g}{ms^2}$)	k_c	$5.0 \cdot 10^{-3}$	$6.0 \cdot 10^{-3}$ (+20%)

Table C.1: Parameter values rescaled to match the fundamental frequency of the recorded coo.

and Herzel (1995). The definition of Q is motivated by the relation $\mathscr{F}_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$ (see Section 2.1.1). Consequently, as explained in Chapter 4 (eq. 4.2 on p. 78), the fundamental frequency of oscillation $\mathscr{F}(t)$ is:

$$\mathscr{F}(t) = Q(t)\mathscr{F}_0 \tag{C.3}$$

According to eq. C.4, the function $Q(t)$ describing the variation of masses and stiffness can be written as:

$$Q(t) \simeq P_{sum}(t)\left(\frac{\max Q - \min Q}{20}\right) + \min Q \tag{C.4}$$

Matching values of $\min Q$ and $\max Q$ with the standard set of parameters (Table 3.2) were found to be in the order of $[\min Q, \max Q] \simeq [0.25, 1]$ (see Fig. C.4). Although recorded and simulated spectrograms show a consistent similarity, we found two inconsistency points with the choice of this range:

① muscle activity *lowers* the default values of tissue and masses ($\min Q, \max Q \leq 1$). It would be more reasonable to hypothesize that a muscle activity modulates the elastic properties of the tissue around the default values. Moreover,

② the bounds of $Q(t)$ ($\min Q, \max Q \leq 1$) cause the syrinx to open in between each trill element, because of the decreased masses and stiffness values which oppose less resistance to the subsyringeal pressure P_s. Therefore, the syrinx closes for constant $Q = 1$ due to the action of P_{TL} as described in Section C.1.1, but opens when we further implement $Q(t)$ as a function of P_{sum}, due to the decreased values of masses and stiffness.

We assumed that $P_{sum}(t)$ modulates tissue mass and stiffness around their default values, and therefore we rescaled the model parameters as shown in Table C.1. The new parameters values rescale the fundamental frequency from $\mathscr{F}_0 \simeq 0.7118$ kHz to $\mathscr{F}_0 \simeq 0.5725$ kHz. In the standard trapezoidal model (Zaccarelli et al., 2006, see Chapter 3), the damping ratio ($\zeta = \frac{r}{2\sqrt{km}}$) - a measure for viscous loss of energy in the focal fold tissue - is in the order of values set for human vocal cords (Ishizaka and Flanagan,

1972; Steinecke and Herzel, 1995). To maintain approximately the same ζ value as in the previous parameters set, we slightly increased r. Consequently, ζ decreases from 0.1118 to 0.0981 ($\simeq -12\%$), but is still in the order of 0.1 (Ishizaka and Flanagan, 1972).

With the new parameter set, the frequency ranges around the default values. We chose the values $[\min Q, \max Q] = [0.8, 1.2]$, i.e., the frequency ranges during coo from $0.8\mathscr{F}_0$ to $1.2\mathscr{F}_0$. The syrinx is closed in between trills and the frequency matches the radiated sound frequency, as described in Chapter 4 (see Fig. C.5).

C.2 Possible improvements

We have shown the process of scaling P_{TL} and P_{sum} into time-series describing variation of a_{0i} and Q, respectively. The spectrogram obtained with the ranges $[\min a_{0i}, \max a_{0i}] = [-3, 1]$ mm^2 and $[\min Q, \max Q] = [0.8, 1.2]$ are qualitatively comparable to the recorded spectrogram of the ring dove coo (see Chapter 4). The syrinx is closed in between trill elements, consistent with experimental observations due to the action of TL. P_{sum} acts as a frequency modulator, varying the parameter Q (and consequently, the elastic properties of the labia, e.g., masses and stiffness) around the default parameters.

The fundamental frequency of the simulated trill matches the fundamental frequency of the recorded coo consistently (see Chapter 4). However, the transmural pressure P_t is very small compared to the pressure P_{TL} that can be exerted on the LVM by the TL muscle (1.0 kPa vs 20 kPa): the resulting frequency range of the simulated trill elements should therefore be increased (compare Fig. C.5 bottom panel with Fig. C.5 top panel, last three trill elements, second and third harmonic). Simulations suggest that the contribution of the transmural pressure P_t in the function P_{sum} is somehow higher than obtained with a simple non-weighted algebraic sum, $P_{sum} = P_{\mathrm{TL}} + P_t$. To maintain closure of the syrinx in between each trill element, we suggest possible improvements in future studies:

- An algebraic sum $P_{sum} = P_{\mathrm{TL}} + \alpha P_t$, $\alpha > 1$. This implementation is equivalent to choose a TL component along the masses displacement, because the TL does not insert perpendicular on the syrinx (Fig. 4.1). Because the functions driving the parameters are scaled, a decreased contribution of TL ($P_{sum} = \beta P_{\mathrm{TL}} + P_t$, $\beta < 1$) corresponds mathematically to an increased contribution of P_t ($P_{sum} = P_{\mathrm{TL}} + \frac{1}{\beta} P_t$,

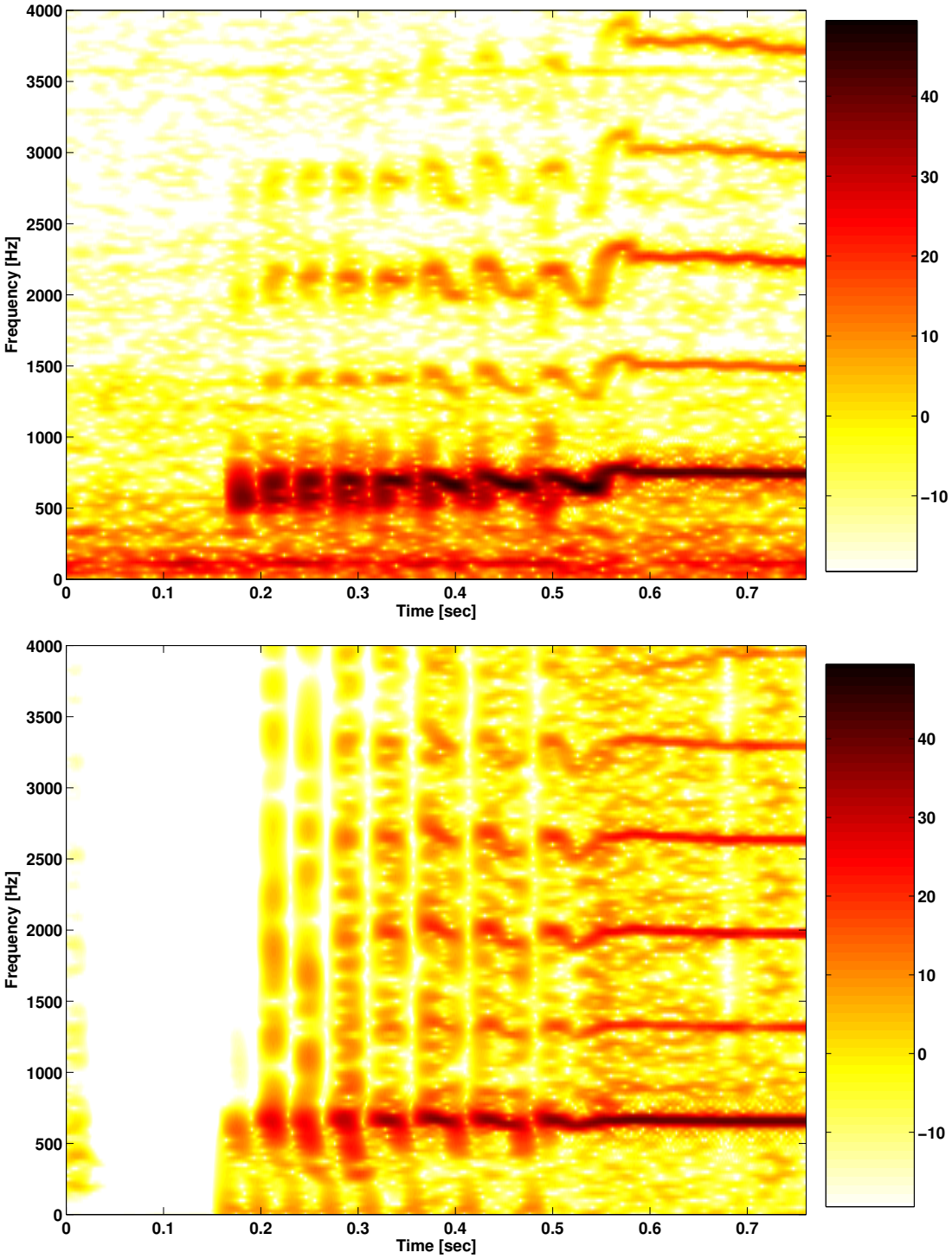


Figure C.4: Spectrograms of a trill. Top panel: recorded ring dove coo. Bottom panel: flow derivative $\dot{U}$ of the simulation obtained with $[\min Q, \max Q] \simeq [0.25, 0.875]$.

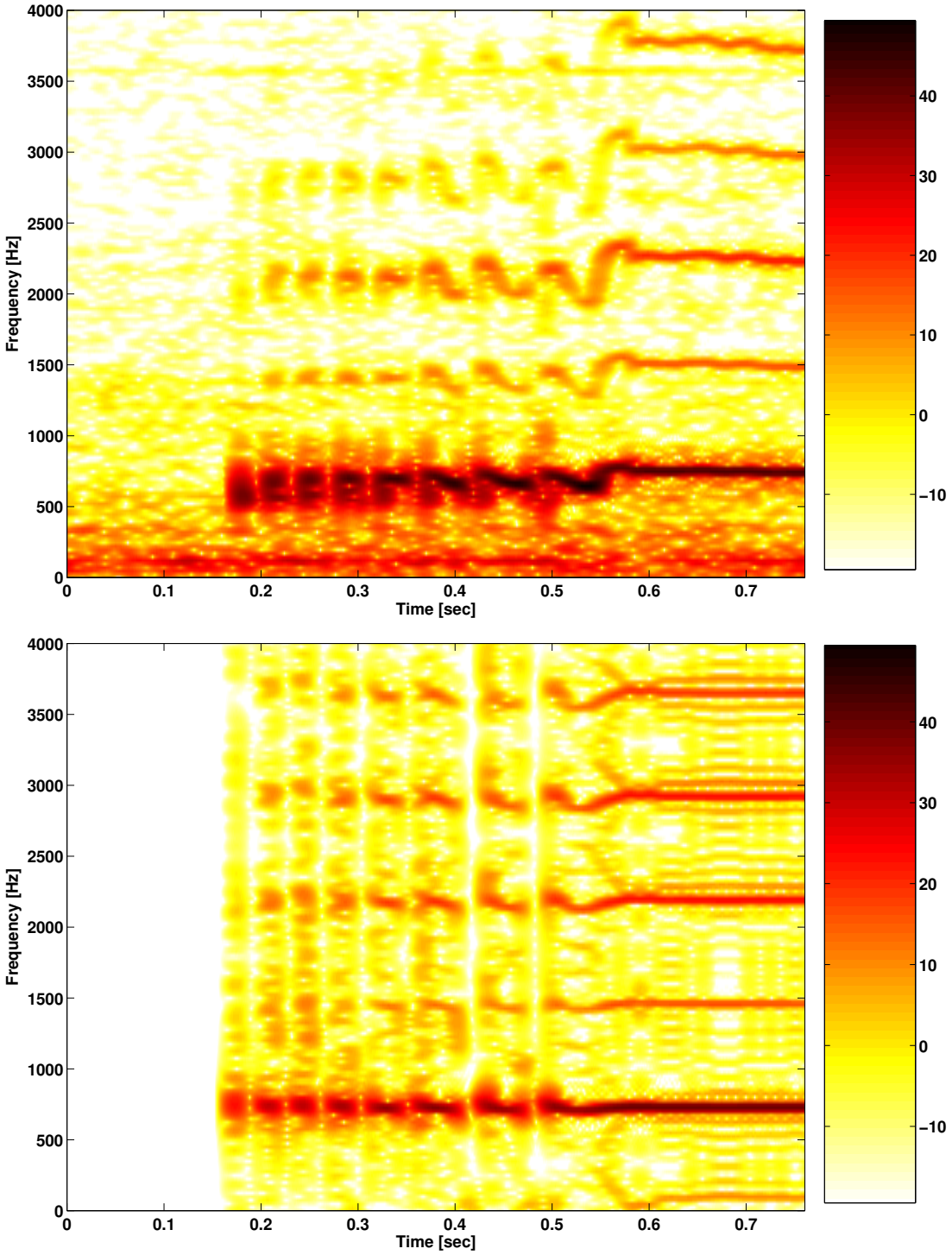


Figure C.5: Spectrograms of a trill. Top panel: recorded ring dove coo. Bottom panel: flow derivative $\dot{U}$ of the simulation obtained with $[\min Q, \max Q] \simeq [0.8, 1.2]$ and the new set of parameters of Table C.1.

$\frac{1}{\beta} > 1$)[1].

- Increase ΔQ, i.e., scale $[\min Q, \max Q]$ to values lower and higher, respectively, than 0.8 and 1.2. This is somehow what has been done with the previous set of parameters (Fig. C.4). However, the scaling function would apply to the function $Q(t) = P_{sum}(t)$, and therefore the decreased value of $\min Q$ would cause again the syrinx to open in between each trill element. Possible implementations of a further driving parameter which controls the minima of Q could be investigated, and could explain the increasing onset of pressure during trill (C. P. H. Elemans, unpublished results). In this framework, we could hypothesize that the ST, which does not affect the output spectrogram and was not implemented for simplicity, could have the role of controlling the position of the syrinx, thus affecting the elastic properties of the tissue and counteracting the decreased $\min Q$.

[1]From equation C.1, assuming $P_{sum} = \frac{1}{\alpha} P_{\mathrm{TL}} + P_t$ and $\min P_{sum} = 0$, the function Q is obtained by means of the relation:

$$\frac{Q - \min Q}{\max Q - \min Q} = \frac{\beta P_{\mathrm{TL}} + P_t}{\max(\beta P_{\mathrm{TL}} + P_t)} = \frac{\beta(P_{\mathrm{TL}} + \frac{1}{\beta} P_t)}{\beta \max(P_{\mathrm{TL}} + \frac{1}{\beta} P_t)} = \frac{(P_{\mathrm{TL}} + \frac{1}{\beta} P_t)}{\max(P_{\mathrm{TL}} + \frac{1}{\beta} P_t)}.$$

Acknowledgments

I am grateful to Hanspeter Herzel, Coen Elemans, Haralambos Hatzikirou, Jürgen Neubauer, Tecumseh Fitch, Tobias Riede and Isao Tokuda for their contribution to this work.

Un grazie sentito a Luca, Lucia e Peggy per avermi accompagnato in questa esperienza.

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfaßt und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Riccardo Zaccarelli
1. Oktober 2007